ARMY RESEARCH LABORATORY

Speech-Based Robotic Control for Dismounted Soldiers: Evaluation of Visual Display Options

by Rodger A. Pettitt, Christian B. Carstens, and Linda R. Elliott

ARL-TR-6946 **May 2014**

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer’s or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5425

ARL-TR-6946 **May 2014**

Speech-Based Robotic Control for Dismounted Soldiers: Evaluation of Visual Display Options

Rodger A. Pettitt, Christian B. Carstens, and Linda R. Elliott
Human Research and Engineering Directorate, ARL

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
May 2014	Final	October 2013–October 2014

4. TITLE AND SUBTITLE
Speech-Based Robotic Control for Dismounted Soldiers: Evaluation of Visual Display Options

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Rodger A. Pettitt, Christian B. Carstens, and Linda R. Elliott

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Laboratory
ATTN: RDRL-HRM-DW
Aberdeen Proving Ground, MD 21005-5425

8. PERFORMING ORGANIZATION REPORT NUMBER
ARL-TR-6946

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Twenty-one enlisted Soldiers supported this user-based evaluation of robot controller options: speech control versus a manual wrist-worn mouse, crossed with two types of visual display options—grid versus object segmentation (OS). Performance on the maneuver course showed slower times for the speech controller, across both levels of workload. The slower times for the speech control can be attributed to the percent of times that the speech control had to be repeated. In addition, the two visual display options were evaluated using a search lane that was cluttered with objects. In this event, the grid display was significantly faster. The OS display was much slower when used with the speech controller. Driving errors were higher in the OS-speech controller condition. Soldier feedback with regard to user experience showed positive regard for the concept of speech control, particularly for its contribution to hands-free operation and maintenance of situation awareness. Improvements for speech control were suggested with regard to robot responsiveness, robot maneuvering, and staying within course limits. The OS display also needs improvement with regard to object labeling updates. As it is currently configured, the labels of the objects can change while the robot is moving, particularly in a cluttered environment.

15. SUBJECT TERMS
robot control, voice control, speech control, Soldier performance, robotic displays

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Rodger A. Pettitt
Unclassified	Unclassified	Unclassified	UU	62	19b. TELEPHONE NUMBER (*Include area code*) 706-545-9142

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Acknowledgments

The authors would like to thank and acknowledge the contributions of project managers Michael Barnes and Sue Hill of the U.S. Army Research Laboratory's Human Research and Engineering Directorate, with reference to multiyear investigations of human-robot interaction. We would also like to thank Jonathan Brown and Chris Blanco from Think-A-Move and Jeff Ferrin from Autonomous Solutions Inc. for providing their insights, equipment, and technical support. We also thank all the support staff that collected performance data alongside each Soldier, and those who diligently organized and reduced the data. We are particularly indebted to the Soldiers who participated in this event. Their expert perspective provided great insight regarding how to best use and improve the advanced controls and displays discussed in this report.

INTENTIONALLY LEFT BLANK.

1. Introduction

1.1 Background

The use of robots in military operations has become pervasive in situations such as ordnance disposal, mine clearing, and remote reconnaissance (Axe, 2008). For these missions, the most common paradigm is one of effortful teleoperation between one user and one robot. Many advancements have been made to overcome performance issues with regard to robot teleoperation to improve processes such as robot responsiveness and camera video bandwidth (Chen et al., 2007). In addition, further advances in technology and artificial intelligence bring more autonomous capabilities, including enhanced perception and object recognition to situation assessment and decisionmaking (Barnes et al., 2014; Schuster et al., 2013), multi-robot multi-operator scenarios (Chen and Barnes, in press; Fincannon et al., 2011), individual differences (Chen, 2011), human-robot trust issues (Hancock et al., 2011), supervisory control (Chen and Barnes, 2012; Chen et al., 2011), and multimodal/bidirectional communications (Lackey et al., 2011), to better support more autonomous robots. As robots become more autonomous, their participation in combat situations expands. Brown (2011) describes operational scenarios that include autonomous robot capabilities as they would impact and support the infantry squad as a decisive force in mission operations. These would include operations such as room clearing, force protection, reducing Soldier load, and recovery/transport of wounded Soldiers. This vision was reiterated by General Robert W. Cone, head of the U.S. Army's Training and Doctrine Command, as the means to accomplish missions while reducing Soldier manpower (Ackerman, 2014).

These advancements present additional challenges and issues for the designer and researcher to identify how best to incorporate robot capabilities to mission requirements (Redden and Elliott, 2010). Given more autonomous robots, the challenge is to identify the most natural means of communicating and controlling robot actions. As robot actions become more tactical and speech recognition algorithms improve, speech-based dialog is expected to be more commonly used in Soldier-robot tasks. Several prototypes have been reported regarding speech control for different military robot applications. Kennedy et al. (2007), using ViaVoice, programmed speech and gesture commands for a core set of commands relevant to U. S. Marine Corps reconnaissance missions: "Attention," "Stop," "Assemble" (i.e., come here), "As you were" (i.e., continue), and "Report" (e.g., assuming robot can communicate to user). Perzanowski et al. (2000a, 2000b, 2002, 2003) reported progress toward a multimodal approach to control single and multiple robots using gesture, smarphone, and speech. Syntactic and semantic information was drawn using ViaVoice and a natural language understanding system (i.e., Nautilus). Jones (2007) reported development of a prototype system capable of detecting and following a person through indoor and outdoor environments while responding to voice and gesture commands. Haas et al.

(2011) investigated integration of speech and touch commands to control a simulation of robot swarms using a visual display. Speech was also demonstrated to be effective in noisy outdoor environments (i.e., outdoor military supply depot) using a "push to talk" approach (Chuangsuwanich et al., 2010). Other approaches to the problem of noisy environments have also been explored (Martinson and Brock, 2007). Speech commands have also been investigated for use within vehicles. For example, Neely et al. (2004) investigated the use of speech commands for robot control when the operator is inside serving as a vehicle commander.

It is clear that while speech controls have been shown to be effective in a variety of military settings, each situation has different issues to be addressed. For example, a system that works well inside a vehicle, integrated with vehicle maps and sound system, may not generalize for use outside the vehicle. A control system that is effective when the operator is sitting still may not necessarily work well when the operator must stand or be on the move. Therefore, there is a need to investigate the factors that are particularly relevant to robot applications for dismounted Soldier use.

This is the third in a series of experiments to investigate advanced controls and displays for robot control by dismounted Soldiers. The first controller experiment focused on aspects of controller size, bulk, and weight (e.g., reducing the size of individual controls and reducing the number of controls through multifunction mapping) (Pettitt et al., 2008). In that study, Soldiers used three types of controllers to maneuver a small unmanned ground vehicle (SUGV) through a driving course that required them to perform various reconnaissance and driving tasks similar to tasks they would normally perform during tactical operations. Several tasks required the operator to control multiple robotic functions simultaneously (e.g., raising the control arm while tilting the camera head). Speech control was expected to reduce mission completion times by enhancing the operator's ability to multitask. For example, the operator could drive the SUGV with the operator control unit (OCU) while using speech commands to pan and zoom a camera or switch camera views. However, results revealed some issues when using speech for basic robot teleoperation tasks. For example, continuous tasks, such as pan, turn, etc., were more difficult to perform with speech control. Because the command was associated with a preset definition, large turns required the operator to give the order multiple times (i.e., saying "turn" once resulted in a 15° turn, but if a 30° turn was needed, the command had to be spoken twice). In addition, it was impossible to fine-tune the continuous commands because of the preset sizes (e.g., the closest the operator could come to a 25° turn was 30°). Recommendations were generated with regard to improvements for integration of speech for robot teleoperation.

In the second study, we focused on the speech commands being used. Soldiers were asked to generate the speech commands they would use for a variety of tasks associated with different robotic reconnaissance missions (Redden et al., 2010). This experiment demonstrated how important it is to tailor speech commands to the target audience. Before training, less than 10% of the commands the Soldiers thought should be used were the commands that were programmed into the speech-control system. Even after training and using many of the commands during a simulation task, only 34% of the Soldiers remembered the commands that the system designers programmed. Commands that were initially intuitive ("Take picture" and "Label alpha") were correctly used by 72% and 83% percent, respectively, of the Soldiers after training. On the other hand, less-intuitive phrases such as "Activate exclusion zone" were not remembered by any of the Soldiers even after training. Thus, it is important to identify and develop intuitive command phrases that are based on familiar military phrases to enhance ease of use and reduce errors, especially during times of combat stress and high cognitive load.

For the third experiment, we refined our approach to speech control by turning attention to more autonomous robot maneuver capabilities. In the first experiment, speech control was found to be less effective on basic teleoperation (movement) tasks than a joystick. However, direct and effortful teleoperation is becoming less necessary as robots gain the capability for autonomous or semi-autonomous waypoint and/or operator-based (e.g., following) navigation. Thus, speech commands are evolving beyond that of basic movements ("Go forward," "Turn left," etc.) that require full attention by the operator to more-general tactical commands such as "Go to location x." In this evaluation, we investigate the capability for speech-based commands to maneuver a robot to specified locations using two different visual display modes. In one display mode, the information given to the robot is based on objects that are recognized and labeled; in the other, the visual display provides a grid system where the location information is given as two coordinates (e.g., x and y). We also compare the speech command capability with a manual command capability (e.g., mouse capability on the wrist-worn visual display).

1.2 Objective

This experiment was designed to investigate the utility of technological improvements in the Think-A-Move (TAM) speech-based robotic control system. We compared the effects of speech control and manual control on the situation awareness (SA) of Soldiers performing dismounted Soldier movement tasks. We also compared the effects of speech control and manual control on movement through a cluttered environment using different autonomous control techniques.

1.3 Overview of Experiment

The study took take place at Fort Benning, GA. Twenty-one Soldiers from engineer and infantry units participated. The Soldiers were asked to maneuver a SUGV through robotic courses using speech control and manual control with and without secondary tasks, as well as two different movement techniques on the display—movement to a grid or movement to a specific object (object segmentation [OS] mode). Performance measures included the number of driving errors, time between waypoints, task completion time, secondary task response rate, and times that the firing hand was removed from a simulated M4 assault rifle to use the robot controller.

2. Method

2.1 Participants

Twenty-one Soldiers from Fort Benning engineer and infantry units participated in the assessment.

2.1.1 Pretest Orientation

The Soldiers were given an orientation on the purpose of the study and what their participation would involve. They were briefed on the objectives, procedures, and robotic system and told how the results would be used and the benefits the military could expect from this investigation. Any questions the subjects had regarding the study were answered.

2.2 Apparatus and Instruments

2.2.1 iRobot PackBot

The iRobot PackBot SUGV with FasTac (figure 1) is a tactical mobile robot used to gather SA information, often in areas that are inaccessible or too dangerous for humans. Its payload has a rotating pan and tilt head equipped with multiple cameras. The PackBot is operated with a manually controlled hand-held controller/display.

Figure 1. PackBot robot.

2.2.2 Robotic Control System

Figure 2 shows the components of the prototype OCU used for the experiment. The backpacked computer worn by the dismounted Soldiers was an Intel Core i5 with 4-GB memory and a Kutta Technologies Unified Ground Control Station system.

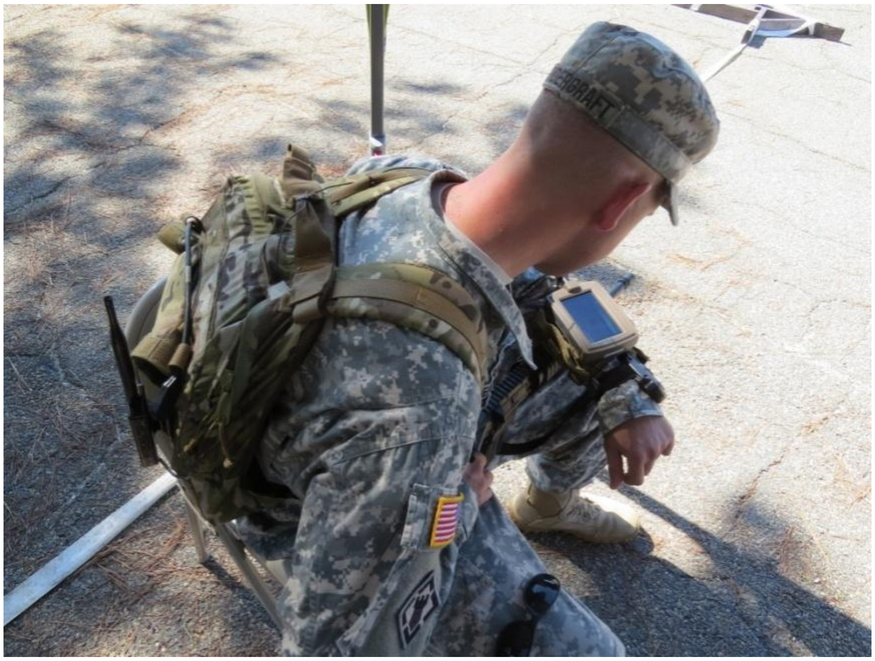

Figure 2. Robotic OCU.

2.2.2.1 SPEAR

The SPEAR (SPEAR Labs, LLC., Nashville, TN) speech control technology used in this experiment was developed by TAM. The SPEAR headset has a 30-dB passive noise cancellation system and an in-ear microphone and speaker. The SPEAR earpiece (figure 3) has a microphone inserted into the ear canal that captures the speech signal. A wired connection carries the signal from the earpiece and ends in a standard 3.5-mm audio jack that can be plugged into a computer soundcard. Through the design of the earpiece, the signal in the desired frequency range can be amplified, thus restoring the quality of the captured speech. A proprietary speech command recognition (SCR) system is used to identify the command spoken by the user. The SCR system collects the captured speech signal and sends out a recognized command. TAM has trained specialized models tuned for recognition of in-ear speech, and the recognition accuracy target has been set to more than 90% even in extreme conditions, which include battlefield conditions where the operator is involved in intense physical activity with loud noise in the background. The Creative (Singapore) Sound Blaster X-Fi Go! high-fidelity USB soundcard allows onboard storage of voice profiles.

Figure 3. SPEAR earpiece.

2.2.3 Video Display

The manipulation software can display information to the operator by using either a grid display (figure 4) or OS display (figure 5). The grid display is overlaid onto the current screen and each grid cell is labeled. This allows the operator to select a grid cell and move to that area. On the OS display, each object is outlined and labeled allowing the operator to select individual objects to move to.

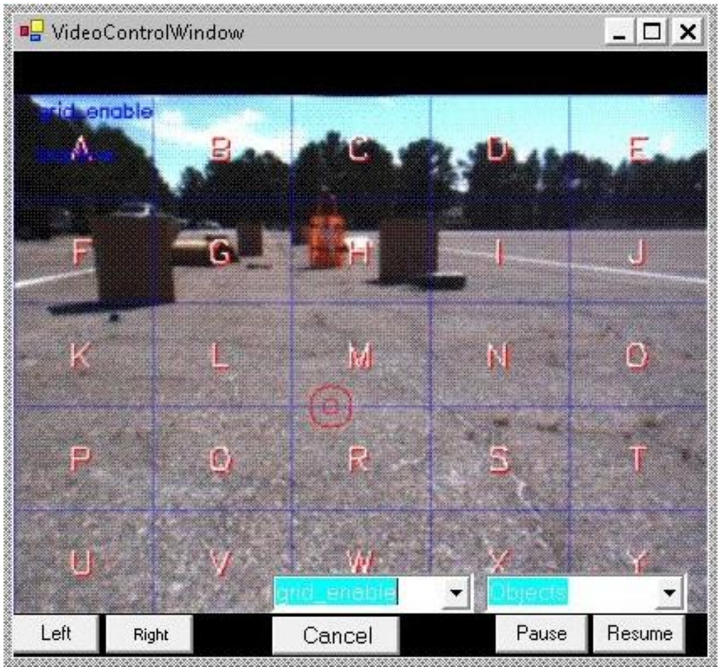

Figure 4. Grid display with labels.

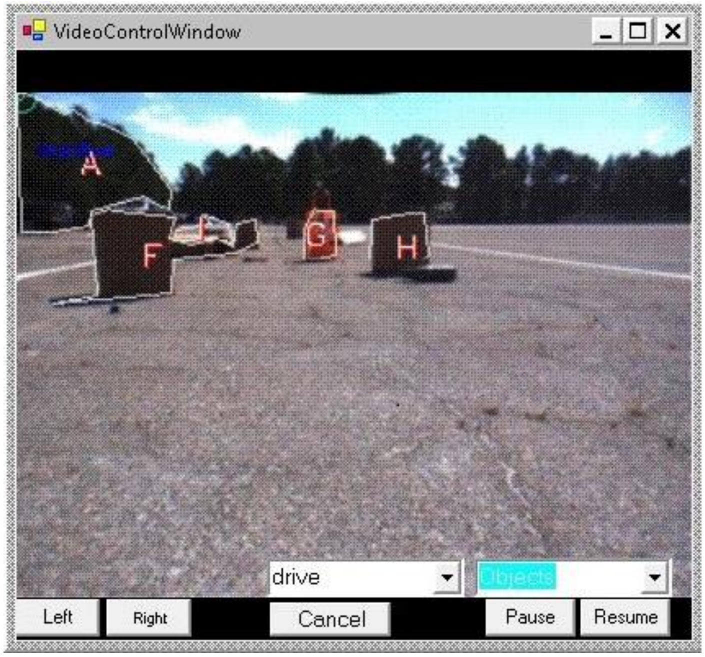

Figure 5. Object segmentation with labels.

During speech-based trials, the operator verbally selected an object or grid based on its label and the robot would autonomously move to the selected location. During manual trials, the operator used a wrist-worn mouse (figure 6) with a joystick-style controller to select a location on the display screen for the robot to move to.

Figure 6. Wrist-worn mouse.

2.2.4 Robotic Course

The robotic course consisted of a maneuver lane and an unexploded ordinance (UXO) search lane.

2.2.4.1 Maneuver Lane

On the maneuver lane (figure 7), Soldiers moved along with the robot in a bounding fashion between waypoints. The lane consisted of three waypoint locations approximately 50 m apart. Once the operator and robot reached the last waypoint, he would then turn the robot around and bound back to the start point, thus completing four bounds. The lane was designed so the operator was facing away from the robot forcing him/her to maneuver the robot using only the driving camera and display. Transition points were marked with red flags along the Soldier's path. The transition points marked the location where the Soldier bounded to prior to maneuvering the robot to the next waypoint.

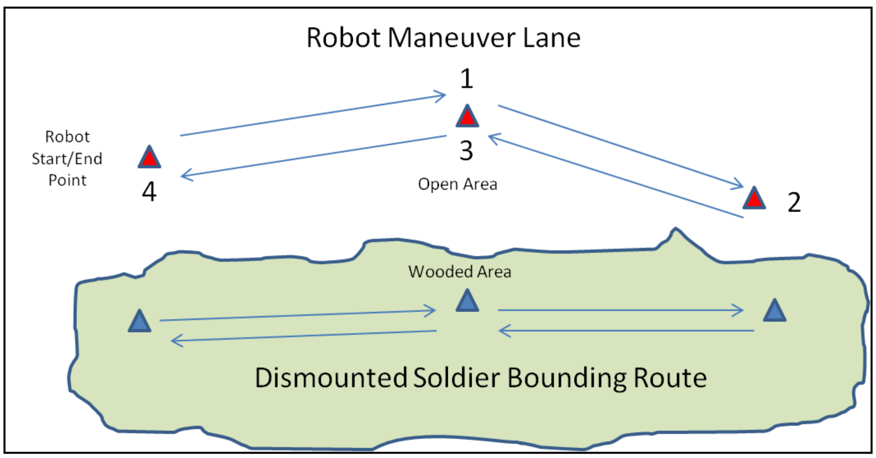

Figure 7. Maneuver lane.

2.2.4.2 UXO Search Lane

The UXO search lane (figure 8) was approximately 100 m long and bordered by engineer tape. Obstacles and inert UXOs were scattered along its length creating a cluttered environment. Soldiers maneuvered the robot through the lane using five marked waypoints as a route guide. The Soldiers operated the robot from a stationary position facing away from the course, forcing them to maneuver the robot using only the driving camera and display.

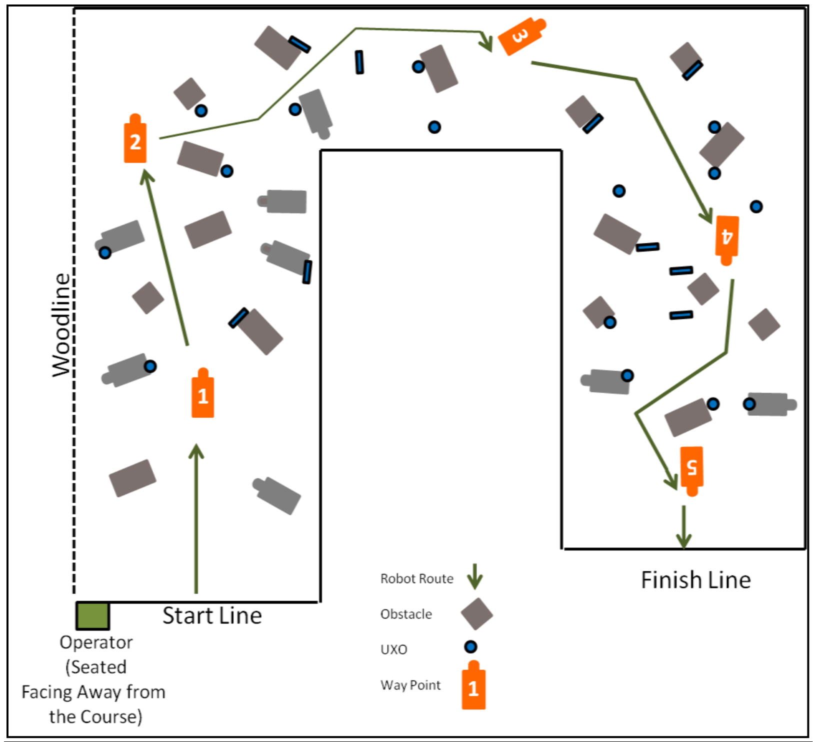

Figure 8. UXO lane.

2.2.5 The National Aeronautics and Space Administration-Task Load Index (NASA-TLX)

The NASA-TLX required the user to rate the workload experienced during task performance on a number of different scales and assign an importance weight to each scale. The scores on the workload scales (mental, physical, temporal, effort, and frustration) can be analyzed separately in weighted or unweighted formats (Hart, 2006; Hart and Staveland, 1988).

2.2.6 Soldier Questionnaires

The questionnaires were designed to elicit Soldiers' opinions about their performance and experiences with each of the controller conditions. The questionnaires asked the Soldiers to provide ratings on a seven-point semantic differential scale ranging from "extremely bad/difficult" to "extremely good/easy." Questionnaires were administered to each Soldier at the end of each iteration (with each control condition) and at the end of the experiment. Questionnaires were also used to gather information concerning the Soldiers' demographic data, robotic experience, and physical characteristics that might affect their ability to operate the robot.

2.3 Procedures

2.3.1 Soldier Orientation

The Soldiers reported in groups of four for one day each, from 0800 to 1700 daily. Upon arrival, they received a roster number used to identify them throughout the evaluation. The Soldiers completed an informed consent form and a demographics questionnaire. They were given an explanation of the robotic mission they would undertake during the experiment. The training and robotic courses were explained, and any questions the Soldiers had concerning the experiment were answered.

2.3.2 Training

A representative from TAM trained the Soldiers on the use of the robot system. Soldiers practiced operating the robot prior to each trial to help mitigate learning effects. They were trained on each control condition just before executing the course with that condition. Soldiers were considered trained once they were able to complete the tasks without assistance. The average training time required was 20 min. Questionnaires concerning the amount of practice time given, the level of detail presented, and the adequacy of training aids were administered at the completion of each training session.

2.3.3 Robotic Course

The robotic course consisted of the maneuver lane and the UXO search lane, described in section 2.2.4. The Soldiers completed two iterations of the robotic course, one using manual control and one using speech control. For each control condition, the maneuver lane was conducted once with secondary tasks and once with no secondary tasks. The UXO search lane was conducted once using the grid display and once using the OS display. Table 1 shows the order of treatment for the maneuver lane and table 2 for the UXO search lane.

Table 1. Maneuver lane order of treatments.

Roster	Maneuver Lane			
	Iteration 1		Iteration 2	
1	Speech/Sec	Speech/NS	Man/Sec	Man/NS
2	Man/Sec	Man/NS	Speech/Sec	Speech/NS
3	Speech/NS	Speech/Sec	Man/Sec	Man/NS
4	Man/NS	Man/Sec	Speech/NS	Speech/Sec
5	Speech/Sec	Speech/NS	Man/NS	Man/Sec
6	Man/Sec	Man/NS	Speech/Sec	Speech/NS
7	Speech/NS	Speech/Sec	Man/NS	Man/Sec
8	Man/NS	Man/Sec	Speech/NS	Speech/Sec
9	Speech/Sec	Speech/NS	Man/Sec	Man/NS
10	Man/Sec	Man/NS	Speech/Sec	Speech/NS
11	Speech/NS	Speech/Sec	Man/Sec	Man/NS
12	Man/NS	Man/Sec	Speech/NS	Speech/Sec
13	Speech/Sec	Speech/NS	Man/NS	Man/Sec
14	Man/Sec	Man/NS	Speech/Sec	Speech/NS
15	Speech/NS	Speech/Sec	Man/NS	Man/Sec
16	Man/NS	Man/Sec	Speech/NS	Speech/Sec
17	Speech/Sec	Speech/NS	Man/Sec	Man/NS
18	Man/Sec	Man/NS	Speech/Sec	Speech/NS
19	Speech/NS	Speech/Sec	Man/Sec	Man/NS
20	Man/NS	Man/Sec	Speech/NS	Speech/Sec
21	Speech/Sec	Speech/NS	Man/NS	Man/Sec
22	Man/Sec	Man/NS	Speech/Sec	Speech/NS
23	Speech/NS	Speech/Sec	Man/NS	Man/Sec
24	Man/NS	Man/Sec	Speech/NS	Speech/Sec

Note: Sec = secondary tasks; NS = no secondary tasks.

Table 2. UXO search lane order of treatments.

Roster	UXO Search Lane			
	Iteration 1		Iteration 2	
1	Speech/Grid	Speech/OS	Man/Grid	Man/OS
2	Man/Grid	Man/OS	Speech/Grid	Speech/OS
3	Speech/OS	Speech/Grid	Man/Grid	Man/OS
4	Man/OS	Man/Grid	Speech/OS	Speech/Grid
5	Speech/Grid	Speech/OS	Man/OS	Man/Grid
6	Man/Grid	Man/OS	Speech/Grid	Speech/OS
7	Speech/OS	Speech/Grid	Man/OS	Man/Grid
8	Man/OS	Man/Grid	Speech/OS	Speech/Grid
9	Speech/Grid	Speech/OS	Man/Grid	Man/OS
10	Man/Grid	Man/OS	Speech/Grid	Speech/OS
11	Speech/OS	Speech/Grid	Man/Grid	Man/OS
12	Man/OS	Man/Grid	Speech/OS	Speech/Grid
13	Speech/Grid	Speech/OS	Man/OS	Man/Grid
14	Man/Grid	Man/OS	Speech/Grid	Speech/OS
15	Speech/OS	Speech/Grid	Man/OS	Man/Grid
16	Man/OS	Man/Grid	Speech/OS	Speech/Grid
17	Speech/Grid	Speech/OS	Man/Grid	Man/OS
18	Man/Grid	Man/OS	Speech/Grid	Speech/OS
19	Speech/OS	Speech/Grid	Man/Grid	Man/OS
20	Man/OS	Man/Grid	Speech/OS	Speech/Grid
21	Speech/Grid	Speech/OS	Man/OS	Man/Grid
22	Man/Grid	Man/OS	Speech/Grid	Speech/OS
23	Speech/OS	Speech/Grid	Man/OS	Man/Grid
24	Man/OS	Man/Grid	Speech/OS	Speech/Grid

Note: OS = object segmentation.

2.3.3.1 Maneuver Lane Procedures

The Soldiers completed the maneuver lane with and without the secondary task for each control condition. Prior to conducting the secondary task trials, the Soldiers were given an operations briefing that explained the robotic mission. During the briefing, it was explained that their primary task was to maneuver the robot along the route as quickly as possible and bound forward each time the robot reached a waypoint. The Soldiers' secondary tasks included monitoring course time using a stop watch, determining number codes using a matrix (table 3), and reporting the activities of a person of interest. The person of interest was located in the vicinity of the operator and performed various scripted activities during the course of the reconnaissance (table 4).

Table 3. Secondary task code matrix.

Task	A	R	B	Q	C	P	D	O	E	N	F	M	G	L	H	K	I	J
1	2	1	0	3	4	6	7	8	9	2	1	0	3	4	6	7	8	9
18	4	3	5	6	8	12	8	9	10	4	3	5	6	8	12	8	9	10
3	6	5	10	9	12	18	9	10	11	6	5	10	9	12	18	9	10	11
16	8	7	15	12	16	24	10	11	12	8	7	15	12	16	24	10	11	12
5	10	9	20	15	20	30	11	12	13	10	9	20	15	20	30	11	12	13
14	12	11	25	18	24	36	12	13	14	12	11	25	18	24	36	12	13	14
7	14	13	30	21	28	42	13	14	15	14	13	30	21	28	42	13	14	15
12	16	15	35	24	32	48	14	15	16	16	15	35	24	32	48	14	15	16
9	18	17	40	27	36	54	15	16	17	18	17	40	27	36	54	15	16	17
10	20	19	45	30	40	60	16	17	18	20	19	45	30	40	60	16	17	18
11	22	21	50	33	44	66	17	18	19	22	21	50	33	44	66	17	18	19
8	24	23	55	36	48	72	18	19	20	24	23	55	36	48	72	18	19	20
13	26	25	60	39	52	78	19	20	21	26	25	60	39	52	78	19	20	21
6	28	27	65	42	56	84	20	21	22	28	27	65	42	56	84	20	21	22
15	30	29	70	45	60	90	21	22	23	30	29	70	45	60	90	21	22	23
4	32	31	75	48	64	96	22	23	24	32	31	75	48	64	96	22	23	24
17	34	33	80	51	68	100	23	24	25	34	33	80	51	68	100	23	24	25
2	36	35	85	54	72	106	24	25	26	36	35	85	54	72	106	24	25	26

Table 4. Secondary task script.

Iteration No.1: Secondary Tasks, 10-s Intervals	
1	Course time
2	Matrix B-5 (20)
3	Person of interest (behind operator with M-4)
4	Matrix P-6 (84)
5	Course time
6	Person of interest (left side behind operator with shovel)
7	Course time
8	Matrix G-2 (54)
9	Person of interest (right side behind operator with radio)
10	Course time
11	Matrix E-10 (18)
12	Matrix H-9 (54)
13	Person of interest (behind operator writing)
14	Course time
15	Person of interest (walking left to right)
16	Matrix J-2 (26)
17	Person of interest (behind operator in prone position)
18	Course time
19	Course time
20	Matrix E-8 (20)
21	Person of interest (behind operator with engineer tape)
22	Person of interest (left behind operator with M-4)
23	Course time
24	Matrix M-13 (60)
25	Matrix L-11 (44)
26	Course time
27	Person of interest (right behind operator talking on radio)
28	Matrix Q-15 (45)
29	Course time
30	Person of interest (standing with shovel)

Soldiers were required to report their responses to the secondary tasks via radio to the data collector. The secondary tasks served as an additional physical and cognitive load. Secondary task questions were asked at 10-s intervals throughout the trial.

During trials both with and without secondary tasks, the Soldiers carried a replica M4 and were instructed to keep their firing hand on the weapon as much as possible while performing the other tasks. A data collector observed the Soldiers and recorded the number of times and total time they removed their firing hands from the weapon.

After each maneuver lane trial, Soldiers completed a NASA-TLX survey.

2.3.3.2 UXO Search Lane Procedures

The Soldiers conducted two trials of the UXO search lane for each control condition, once using the grid display configuration and once using the OS configuration. The Soldiers were instructed to maneuver the robot through the lane using the marked waypoints as a guide and report any UXO they observed. They were also instructed to stay within the course boundaries marked with white engineer tape and avoid hitting obstacles and UXO with the robot.

Soldiers were given a replica M4 and were instructed to keep their firing hand on the weapon as much as possible while maneuvering the robot through the course. A data collector following the robot recorded course completion time, number of driving errors, and number of times off course. A driving error was recorded when the robot ran into an obstacle or over UXO, and an off course error was recorded if either track of the robot left the boundary of the lane. A data collector co-located with the operators recorded number of UXO reported and the number of times and total time they removed their firing hands from the weapon. The locations of the waypoints were changed between trials to minimize learning effects. After each UXO search lane trial, Soldiers completed a NASA-TLX survey. Upon completing each iteration, the Soldiers were given a questionnaire designed to assess their performance and experiences with each of the control systems.

2.3.4 End-of-Experiment Questionnaire

After completing the manual and speech control iterations, the Soldiers completed an end-of-experiment questionnaire that compared each of the control methods on a number of characteristics.

3. Results

The following section provides descriptive statistics (means, standard deviations) and inferential tests of significance based on repeated measures general linear model (GLM). In addition to the F and p statistics associated with hypothesis testing, and df (degrees of freedom), we also provide a measure of effect size, partial eta square ($\eta\rho^2$), to better interpret the practical significance of differences between conditions. In addition, we provide visual representation of the 95% confidence intervals within the bar graphs, which indicate the precision of the mean, as it generalizes to the population (i.e., the range estimated, with 95% confidence, to include the “true” mean).

3.1 Demographics

The 21 participants were enlisted Soldiers assigned to Fort Benning, GA, whose military occupational specialties included 5 infantry, 14 engineers, 1 electronic warfare specialist, and 1 Special Forces candidate. Ranks included 10 privates (E1–E3), 6 specialists (E4), and 5 sergeants (E5) with a mean time in service of 32 months (ranging 6–96 months). Four of the Soldiers had military experience in teleoperating ground unmanned robots. Detailed responses to the demographics questionnaire are available in the appendix.

3.2 Training

Participants stated that the training was thorough and fully prepared them to perform the tasks required to conduct the robotic reconnaissance course. "Learning to operate the controls and drive the robot" was rated as being very easy for all control conditions. Several Soldiers indicated that learning to maneuver the robot using the OS display was more difficult than with the grid display because of the object labels changing as the robot moved. During initial iterations, several Soldiers experienced difficulty stopping the robot in both the manual and speech modes. The robot's response to the stop command improved after software changes were applied. Detailed responses to the training questionnaire are available in the appendix.

3.3 Maneuver Course

The mean course completion times for the maneuver course are shown in table 5 and figure 9. A repeated measures analysis of variance (ANOVA) showed that the mean completion time for speech control was significantly slower than the mean time for manual control, $F(1,20) = 16.15$, $p = .001$, $\eta^2_p = .447$. The secondary tasks were effective in distracting the participants from operating the robot; mean course completion time was significantly slower with the addition of the distracter tasks, $F(1,20) = 23.08$, $p < .001$, $\eta^2_p = .536$. There was no significant Control x secondary task interaction, $F < 1$.

Table 5. Mean and (standard deviation) course completion times, maneuver course.

Secondary Tasks	Manual (s)	Speech (s)
No tasks	308.9 (35.3)	336.3 (34.0)
Tasks	355.7 (42.8)	383.8 (51.1)

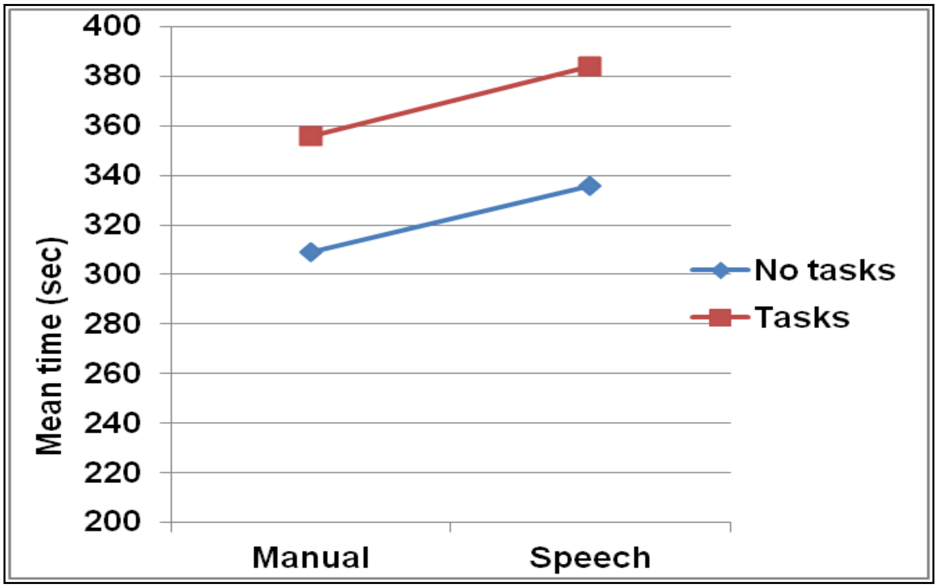

Figure 9. Mean course completion times, maneuver course.

The slower times can be attributed to the number of times the speech commands had to be repeated before the robot responded. The number of times that the participants had to repeat speech commands in order to be understood by the robot was recorded. There were three command verbs: “Select,” “Spin,” and “Stop.” Table 6 shows the percentage of times each command had to be repeated. The differences in percent repeats for the three commands approached statistical significance, $F(2,40) = 3.11$, $p = .055$, $\eta^2_p = .135$.

Table 6. Percent repeated commands.

Command	Mean	Statistical Difference
Select	17%	11%
Spin	31%	21%
Stop	29%	24%
All	26%	20%

Summary statistics for the NASA-TLX workload measure on the maneuver course are shown in table 7. The repeated measures ANOVAs for the NASA-TLX scale scores are summarized in table 8. For each of the scales, the main effects for tasks (i.e., with versus without secondary tasks) were statistically significant, while the *F* values for Control and for the Control x Tasks interaction were insignificant. The main effects for tasks are shown in figure 10. The secondary tasks clearly increased the perceived workload in both the speech and manual control conditions, while the controller type had no significant impact on workload.

Table 7. Means and (standard deviations), maneuver course.

Scale	Manual, No tasks (s)	Manual, Tasks (s)	Speech, No tasks (s)	Speech, Tasks (s)
Mental	2.81 (1.50)	4.67 (1.85)	2.71 (1.42)	4.05 (1.80)
Physical	2.48 (1.69)	3.33 (2.13)	2.38 (1.47)	2.90 (1.41)
Temporal	2.90 (2.05)	4.52 (2.20)	2.57 (1.66)	3.81 (1.89)
Effort	2.67 (1.53)	4.33 (2.20)	2.48 (1.50)	3.81 (1.81)
Frustration	2.29 (1.23)	3.43 (1.99)	2.38 (1.83)	3.10 (1.81)

Table 8. Summary of ANOVAs, NASA-TLX, maneuver course.

Source	Statistic	Mental	Physical	Temporal	Effort	Frustration
Control	F	2.32	2.66	2.27	2.96	<1
	df	1,20	1,20	1,20	1,20	1,20
	p	0.143	0.118	0.147	0.101	—
	η^2_p	0.104	0.116	0.102	0.129	—
Tasks	F	69.8	12.33	54.8	61.0	15.6
	df	1,20	1,20	1,20	1,20	1,20
	p	<0.001*	0.002*	<0.001*	<0.001*	0.001*
	η^2_p	0.777	0.381	0.733	0.753	0.439
Control x Tasks	F	2.79	<1	<1	<1	1.26
	df	1,20	1,20	1,20	1,20	1,20
	p	0.11	—	—	—	0.275
	η^2_p	0.122	—	—	—	0.059

* $p < .05$

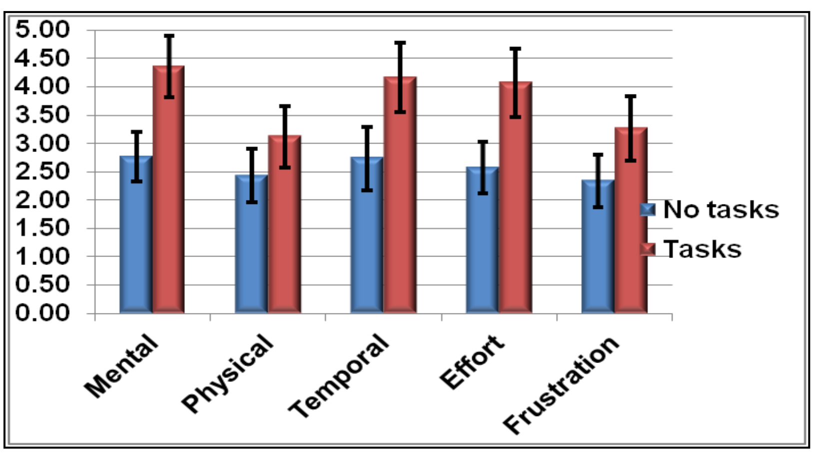

Figure 10. Mean NASA-TLX scale scores with 95% confidence interval, maneuver course.

Figure 11 shows the mean number of times that the participants removed their firing hand from the simulated M4 while driving the robot through the maneuver course. There was a significant main effect for Control, $F(1,20) = 100.0$, $p < .001$, $\eta^2_p = .833$, and for tasks, $F(1,20) = 27.0$, $p < .001$, $\eta^2_p = .574$. While the Soldiers did not have to use their firing hand (as opposed to their other hand) to accomplish the secondary task, it can be seen that some Soldiers did. There was no significant Control x Task interaction: $F < 1$.

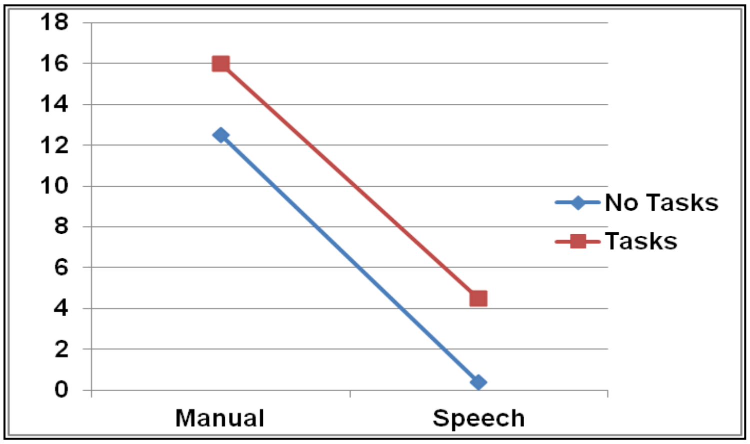

Figure 11. Mean number of times firing hand removed from weapon, maneuver course.

Participants answered the secondary task questions correctly 76% of the time when using manual control and 78% of the time with speech control. The difference was not statistically significant: $t < 1$.

3.4 UXO Course

The mean UXO course completion times are shown in figure 12. The main effect for Control approached statistical significance, $F(1,20) = 4.20$, $p = .054$, $\eta^2_p = .174$, with manual control being somewhat faster than speech control. The main effect for Display was statistically significant, $F(1,20) = 7.48$, $p = .013$, $\eta^2_p = .272$, with the grid display being on average faster than the OS display. The interaction Control x Display, $F(1,20) = 5.83$, $p = .025$, $\eta^2_p = .226$, was also significant. Post hoc paired comparisons (table 9) indicate that course times for the two display modalities were not significantly different using manual control but that the grid display produced significantly faster times with speech control. Times were slowest for the combination of speech and OS display.

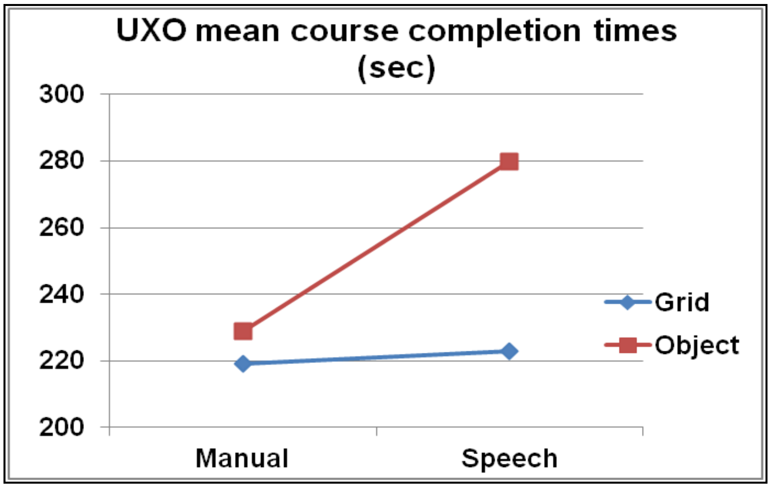

Figure 12. Mean course completion times, UXO course.

Table 9. Post hoc paired comparisons, UXO course completion times.

Mean 1	**Mean 2**	***t***
Manual-grid	Manual-object	<1
Speech-grid	Speech-object	4.17*
Manual-grid	Speech-grid	<1
Manual-object	Speech-object	3.69*

*p <.05, 2-tailed

Summary statistics for the NASA-TLX workload measure on all equipment conditions in the UXO course are shown in table 10.

Table 10. Means and (standard deviations), NASA-TLX, UXO course.

Scale	**Manual Grid (s)**	**Manual OS (s)**	**Speech Grid (s)**	**Speech OS (s)**
Mental	4.10 (2.36)	4.19 (2.52)	3.71 (2.12)	4.24 (2.26)
Physical	2.14 (1.31)	2.24 (1.41)	2.06 (1.60)	2.05 (1.53)
Temporal	3.43 (1.94)	3.48 (2.23)	3.52 (1.78)	3.62 (1.88)
Effort	3.76 (2.55)	3.62 (1.91)	3.24 (1.58)	4.05 (2.27
Frustration	3.19 (2.27)	3.19 (2.11)	3.05 (1.75)	3.90 (2.45)

ANOVAs on the NASA-TLX scale scores are summarized in table 11. None of the main effects or interactions was statistically significant. However, given the consistent findings with regard to Soldier experience with the OS display (i.e., more difficult), planned comparison analyses were applied to the results regarding TLX measures of mental workload, effort, and frustration with regard to visual grid and OS displays (see table 12). Results from this analysis show significantly higher levels of frustration associated with the OS display.

Table 11. Summary of ANOVAs, UXO course.

Source	Stat	Mental	Physical	Temporal	Effort	Frustration
Control	*F*	<1	<1	<1	<1	<1
	df	1,20	1,20	1,20	1,20	1,20
	p	—	—	—	—	—
	η^2_p	—	—	—	—	—
Display	*F*	<1	<1	<1	1.47	2.95
	df	1,20	1,20	1,20	1,20	1,20
	p	—	—	—	0.239	0.101
	η^2_p	—	—	—	0.069	0.129
Control x Display	*F*	<1	<1	<1	3.50	2.95
	df	1,20	1,20	1,20	1,20	1,20
	p	—	—	—	0.076	0.101
	η^2_p	—	—	—	0.149	0.129

Table 12. Direct planned comparison results for TLX scales regarding speech grid vs. speech OS.

Scale	*t*	df	*p*
Mental	1.33	20	0.199
Effort	2.02	20	0.057
Frustration	2.12	20	0.047

Summary statistics for driving errors, number of times off course, and number of UXOs detected are shown in table 13. For the off-course errors, neither the main effects for Control, $F < 1$, nor Display, $F(1,20) = 2.91$, $p = .104$, $\eta^2_p = .127$, were statistically significant. The Control x Display interaction was not statistically significant. There were no significant effects for number of UXOs detected: all $F < 1$. For driving errors, the main effects for Control, $F < 1$, and Display, $F(1,20) = 1.30$, $p = .268$, $\eta^2_p = .061$, were not statistically significant considered in the overall ANOVA. However, given the trend with regard to the grid and OS displays, as reflected in the outcome measures for driving errors, a planned comparison analysis was also performed. This more powerful and specific comparison showed a significant difference, with OS associated with higher average driving errors ($t = 3.25$, df = 20, $p = .004$).

Table 13. Means (standard deviations), UXO course.

Measure	Manual Grid (s)	Manual OS (s)	Speech Grid (s)	Speech OS (s)
Driving errors	1.10 (1.18)	0.76 (0.72)	0.43 (0.68)	1.14 (0.96)
Off course	0.33 (0.73)	0.62 (1.02)	0.43 (0.60)	0.52 (0.75)
UXOs	10.95 (3.92)	10.71 (3.36)	10.67 (3.75)	10.81 (3.82)

The Control x Display interaction, shown in figure 13, was statistically significant: $F(1,20) = 6.15$, $p = .022$, $\eta^2_p = .235$. However, post hoc paired using the Bonferroni correction for family-wise error (table 14) indicates that none of the paired comparisons were statistically significant. There was a trend for fewer driving errors with the grid display when the participants used speech control but no advantage for the grid display with manual control.

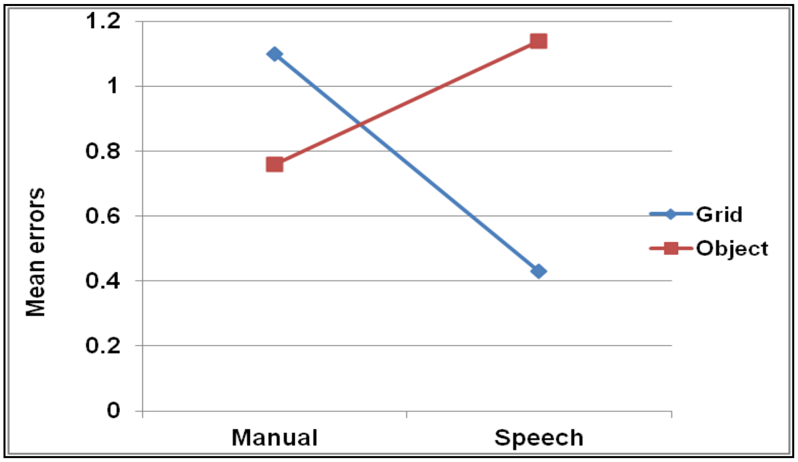

Figure 13. Mean number of driving errors, UXO course.

Table 14. Post hoc paired comparisons, UXO driving errors.

Mean 1	Mean 2	*t*
Manual-grid	Manual-OS	1.14
Speech-grid	Speech-OS	2.38
Manual-grid	Speech-grid	2.24
Manual-OS	Speech-OS	1.27

Figure 14 shows the mean number of times that the participants removed their firing hand from the weapon while driving the robot through the UXO course. There was a significant main effect for Control, $F(1,20) = 82.0$, $p < .001$, $\eta^2_p = .804$, showing a much higher rate of hand removal in the manual condition. Neither the main effect for Display ($F < 1$) nor the Control x Display interaction ($F < 1$) were statistically significant.

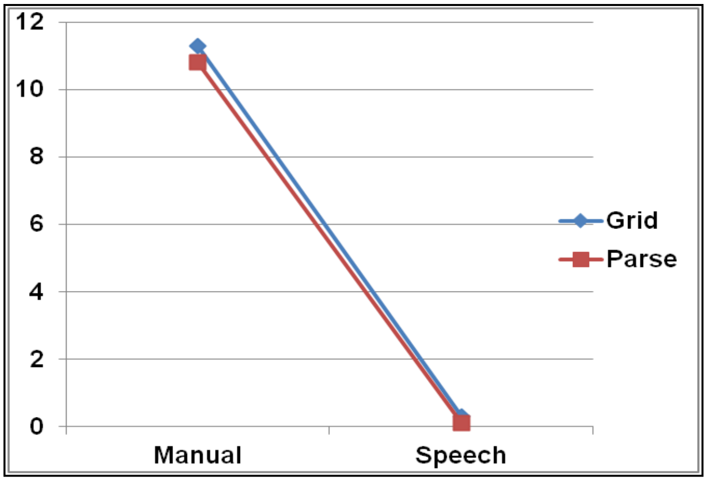

Figure 14. Mean number of times firing hand removed from weapon, UXO course.

3.5 End-of-Experiment Questionnaire

After completing the manual and speech control iterations, the Soldiers were asked to select the control condition they felt was most effective for performing robotic control tasks on each of the courses. Figure 15 shows the Soldiers' preference between the speech and manual control conditions when using the grid display for each task listed. Figure 16 shows the Soldiers preference between the speech and manual control conditions when using the OS display for each task listed. The majority of Soldiers preferred speech over manual for secondary task performance and for maintaining SA while maneuvering the robot. Also, the hands-free nature of speech control allowed the Soldiers to keep their hands on their weapons as opposed to manipulating a control device. Twenty of the 21 participants stated that speech-based control shows potential for robotic operation.

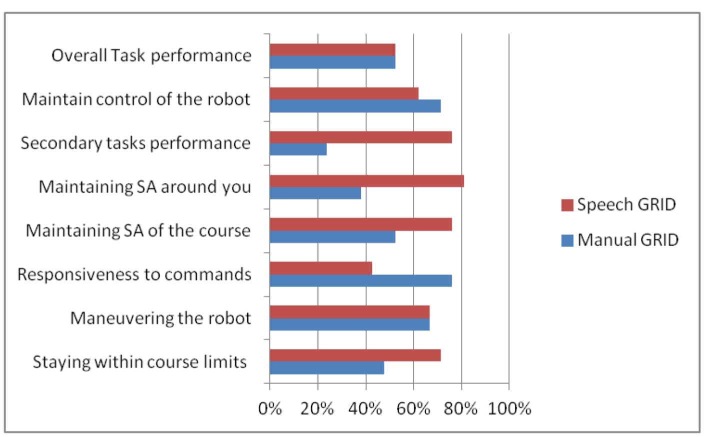

Figure 15. Grid display control condition preference for task performance.

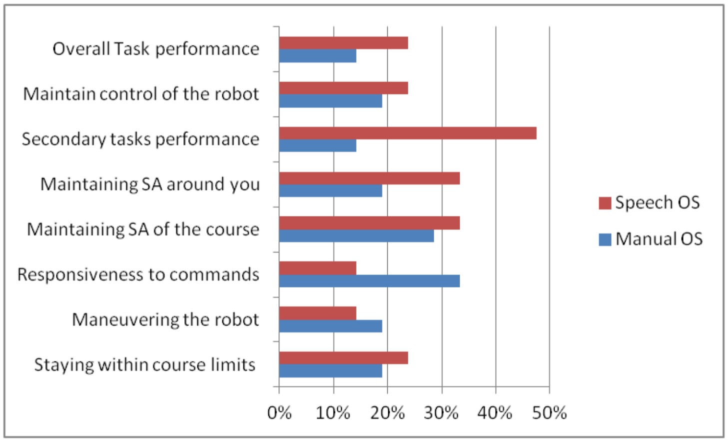

Figure 16. OS display control condition preference for task performance.

Figures 17 and 18 shows that the Soldiers preferred using the grid display over the OS for all tasks. The preference was based on the Soldiers' ability to retask the robot while it was moving. When using the OS, the system continuously relabeled objects while it was moving. Before selecting a different object for the robot to go to, the operator had to first stop the robot and select a new label. When using the grid display, the operator was able to select a new grid while the robot was moving. The ability to retask the robot while it was maneuvering enhanced the operator's ability to maneuver around object and direct the robot along a desired path.

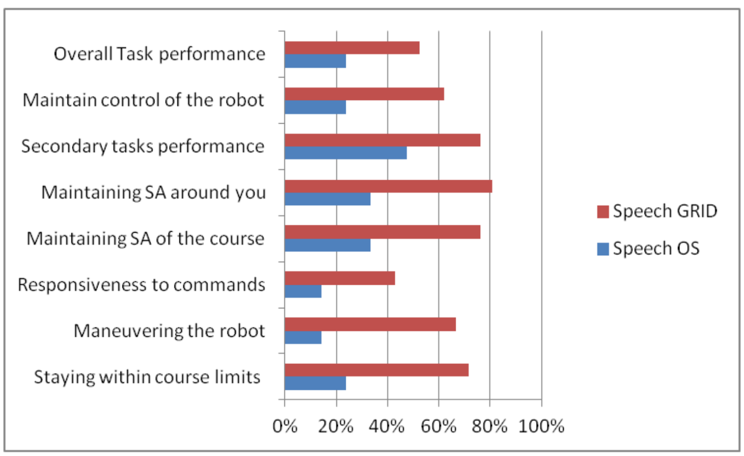

Figure 17. Speech grid vs. speech OS.

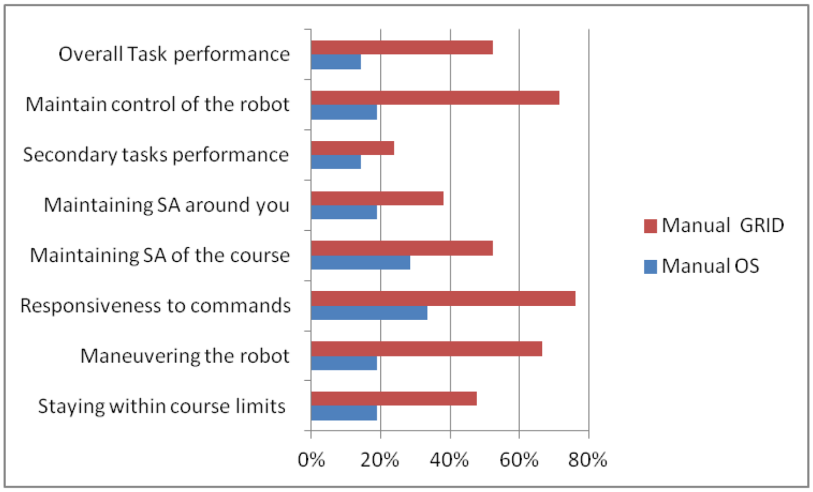

Figure 18. Manual grid vs. manual OS.

Detailed responses to the end of experiment questionnaire are available in the appendix.

4. Discussion and Recommendations

Twenty-one enlisted Soldiers participated in this user-based evaluation. They were trained to use two different controller options, speech and manual (wrist-worn mouse) control, crossed with two types of visual display options, grid and object segmentation, for controlling a PackBot robot. The Soldiers indicated that the training was sufficient and that, in general, it was easy to learn to use each controller and display option. However, there was some difficulty with the OS visual display option because of the object labels changing as the robot moved through a cluttered environment. During initial iterations, several Soldiers experienced difficulty stopping the robot in both the manual and speech modes. The robot's response to the stop command improved after software changes were applied.

There were two main approaches for experiment task performance. The maneuver course was used to assess differences between the two controller options, speech versus manual, under two levels of workload, as manipulated by a secondary task (e.g., locating information on a look-up table). For the maneuver course, the visual display was held constant as the focus was on the navigation waypoint maneuvers with and without secondary tasks. That is, in the maneuver course, we contrasted the two controllers (speech and manual) under relatively low and high workload. In contrast, the UXO search lane, which was cluttered with objects, was used to examine differences between the two visual display options when used with the two different controllers.

Performance on the maneuver course showed slower times for the speech controller across both levels of workload. Workload also had a main effect on task time, i.e., having the secondary task was associated with slower times. As expected, trials having the secondary task were associated with ratings of higher workload. The slower times for the speech control can be attributed to the percent of times that the speech control had to be repeated, which ranged from 17% to 31%.

Comparison of the two visual display options showed a distinct difference with the grid display on average being significantly faster. In addition, the interaction was significant, such that the OS display was much slower when used with the speech controller. Driving errors were also higher in the OS speech controller condition,

The hands-free advantage of speech control was quantified by the number of times (mean = 12) the operator had to remove his/her hand from the weapon to use the manual controller as opposed to the speech controller (mean = 0). In the maneuver course, the speech control was also associated with much lower interference with the Soldiers control of their weapon, though there was some interference due to the secondary tasks. While the Soldiers did not have to use their firing hand as opposed to their other hand to accomplish the secondary tasks, some Soldiers chose to do so.

Soldier feedback showed positive regard for the concept of speech control, particularly for its contribution to hands-free operation and maintenance of SA. However, improvements for speech control were suggested with regard to robot responsiveness, robot maneuvering, and staying within course limits. In addition, ratings for speech control were higher when associated with the visual grid. The OS display needs improvement with regard to object labeling updates. As it is currently configured, the labels of the objects can change while the robot is moving, particularly in a cluttered environment. The preference was based on the Soldiers' ability to retask the robot while it was moving. When using the OS, the system continuously relabeled objects while it was moving. Before selecting a different object for the robot to go to, the operator had first stop the robot and select a new label. When using the grid display, the operator was able to select a new grid while the robot was moving. The ability to retask the robot while it was maneuvering enhanced the operator's ability to maneuver around objects and direct the robot along a desired path.

Thus, core findings from this report include the following:

- Speech control is a promising concept of operations.
- Soldiers appreciate the hands-free aspect of speech control.
- Speech control device was easy to learn.
- Speech control needs improvement to reduce repetitions.

- Grid-based visual display associated with higher performance in cluttered environments.
- Need to refine the OS visual display for ease of use.

Future improvements in robot control would also include integrating speech control with other forms of intuitive control. While this report focused on the integration with visual displays, there is the challenge of making the control eyes-free as well as hands-free. The use of gestures to enhance speech-based controls shows promise regarding communication of spatial relationships and explicit directions (e.g., “Go to the east side of the third building behind the church.”). Pointing gestures are expected to help clarify localization information without the need to manipulate a visual display. The challenge becomes even greater if one is commanding an aerial robot having even more variety of directional commands; neither speech nor any programming language is well suited for these types of commands (Hirzinger, 2001).

Speech-based controls can be problematic in noisy environments such as aircraft carriers. This challenge might also be ameliorated with integration of speech and gesture. Urban and Bajcsy (2005) developed and demonstrated the usefulness of a combined speech and gesture approach to the taxiing commands of unmanned aerial vehicles when landing and taking off in such noisy situations. Iba et al. (2003) reported progress with regard to robot capability to infer user intention from speech and gesture commands. Brooks (2005) reported progress toward naturalistic interaction with robots for Soldier tasks such that speech control was augmented by robot capability to learn and imitate tasks and use of social interaction cues (e.g., gaze direction, nodding, and facial expressions), and codified verbal expressions. Robot capabilities included detection of head orientation, body motion mimicry, hand reflex, figure-ground segmentation, and response to operator gestures and/or touch. Prasov (2012) explored the role of shared gaze to augment language understanding between operator and robot in military operations. Thus, the challenges regarding speech-based control include not only those regarding speech intelligibility (e.g., brevity, clarity, and ease of understanding), and integration with the most appropriate visual displays, but also integration of speech commands with other naturalistic multimodal channels of communication.

5. References

Ackerman, E. U.S. Army Considers Replacing Thousands of Soldiers With Robots. *IEEE Spectrum,* 22 January 2014. http://spectrum.ieee.org/automaton/robotics/military-robots/army-considers-replacing-thousands-of-soldiers-with-robots 2014 (accessed 22 January 2014).

Axe, D. *War Bots: U.S. Military Robots Are Transforming War in Iraq, Afghanistan, and the Future;* Nimble Books LLC: Ann Arbor, MI, 2008.

Barnes, M.; Chen, J.; Jentsch, F.; Oron-Gilad, T.; Redden, E.; Elliott, L.; Evans, A. *Designing for Humans in Autonomous Systems: Military Applications;* ARL-TR-6782; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, January 2014.

Brooks, R. *Natural Tasking of Robots Based on Human Interaction Cues;* Technical Report ADM001779; Massachusetts Institute of Technology, Cambridge Laboratory for Computer Science: Cambridge, MA, 2005.

Brown, R. B. The Infantry Squad: Decisive Force Now and in the Future. *Military Review* **2011,** *91* (6), 2–9.

Chen, J. Individual Differences in Human-Robot Interaction in a Military Multitasking Environment. *Journal of Cognitive Engineering and Decision Making* **2011,** *5* (1), 83–105.

Chen, J.; Barnes, M. Supervisory Control of Multiple Robots: Effects of Imperfect Automation and Individual Differences. *Human Factors* **2012,** *54* (2), 157–174.

Chen, J.; Barnes, M. Human-Agent Teaming For Multi-Robot Control: A Review of Human Factors Issues. *IEEE Transactions on Human-Machine Systems,* in press.

Chen, J.; Barnes, M.; Harper-Sciarini, M. Supervisory Control of Multiple Robots: Human Performance Issues and User Interface Design. *IEEE Transactions on Systems, Man, and Cybernetics: Part C: Applications and Reviews* **2011,** *41* (4), 435–454.

Chen, J.; Haas, E.; Barnes, M. Human Performance Issues and User Interface Design for Teleoperated Robots. *IEEE Transactions on Systems, Man, and Cybernetics: Part C: Applications and Reviews* **2007,** *37* (6), 1231–1245.

Chuangsuwanich, E.; Cyphers, S.; Glass, J.; Teller, S. Spoken Command of Large Mobile Robots In Outdoor Environments. *Proceedings of the 2010 IEEE Spoken Language Technology Workshop,* Berkeley, CA, 12–15 December 2010; pp 306–311.

Fincannon, T.; Keebler, J.; Jentsch, F.; Phillips, E.; Evans, W. Team Size, Team Role, Communication Modality, and Team Coordination in the Distributed Operation of Multiple Heterogeneous Unmanned Vehicles. *Journal of Cognitive Engineering and Decision Making* March **2011,** *5* (1), 106–131.

Haas, E.; Pillalamarri, K.; Stachowiak, C.; Fields, M. Multimodal Controls for Soldier/Swarm Interaction. *Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication*, Atlanta, GA, 31 July–3 August 2011; pp 223–228.

Hancock, P.; Billings, D.; Oleson, K.; Chen, J.; De Visser, E.; Parasuraman, R. *A Meta-Analysis of Factors Influencing the Development of Human-Robot Trust;* ARL-TR-5857; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2011.

Hart, S. NASA-Task Load Index (NASA-TLX): 20 Years Later. *Human Factors and Ergonomics Society Annual Meeting Proceedings* **2006,** *50*, 904–908.

Hart, S.; Staveland, L. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload;* Hancock, P., Meshkati, N., Eds.; Amsterdam: North Holland, 1988; pp 139–183.

Hirzinger, G. On the Interaction Between Human Hand and Robot – From Space to Surgery. *Proceedings of the 2001 IEEE International Workshop on Robot and Human Interactive Communication,* Paris, France, 18–21 September 2001.

Iba, S.; Paredis, C.; Khosla, P. Intention Aware Interactive Multi-Modal Robot Programming. *Proceedings of the 2003 IEEE/RJS International Conference on Intelligent Robots and Systems,* Las Vegas, NV, October 2003.

Jones, C. *Tactical Teams: Cooperative Robot/Human Teams for Tactical Maneuvers;* DARPA Order U051-48; U.S. Army Aviation and Missile Command/Defense Advanced Research Projects Agency, 2007.

Kennedy, W.; Bugajska, M.; Marge, M.; Adams, W.; Fransen, B.; Perzanowski, D.; Schultz, A.; Trafton, G. Spatial Representation and Reasoning for Human-Robot Collaboration. *Proceedings of the 22nd Conference on Artificial Intelligence,* Vancouver, Canada, 22–26 July 2007; AAAI Press: Danvers, MA; pp 1554–1559.

Lackey, S. J.; Barber, D.; Reinerman-Jones, L.; Badler, N.; Hudson, I. Defining Next-Generation Multi-Modal Communication in Human-Robot Interaction. *Proceedings of the Human Factors and Ergonomics Society Conference 2011*, Las Vegas, NV, 19–23 September 2011.

Martinson, E.; Brock, D. Improving Human-Robot Interaction Through Adaptation to the Auditory Scene. *Proceedings of the 2nd ACM/IEEE International Conference on Human Robot Interaction*, Washington, DC, 9–11 March 2007; pp 113–120.

Neely, H.; Belvin, R.; Fox, J.; Daily, M. Multimodal Interaction Techniques for Situational Awareness and Command of Robotic Combat Entities. *Proceedings of the 2004 IEEE Aerospace Conference,* Big Sky, MT, 6–13 March 2004; pp 3297–3305.

Perzanowski, D.; Adams, W.; Schulz, A.; Marsh, E. *Toward a Seamless Integration in a Multi-Modal Interface;* ADA434973; U.S. Navy Center for Applied Research in Artificial Intelligence, U.S. Naval Research Laboratory: Washington, DC, 2000a.

Perzanowski, D.; Schultz, A.; Adams, W.; Marsh, E. *Using a Natural Language and Gesture Interface for Unmanned Vehicles;* ADA435161; U.S. Navy Center for Applied Research in Artificial Intelligence, U.S. Naval Research Laboratory: Washington, DC, 2000b.

Perzanowski, D.; Schultz, A.; Adams, W.; Bugajska, M.; Marsh, E.; Trafton, G.; Brock, D. *Communicating With Teams of Cooperative Robots;* unnumbered technical report; Naval Research Laboratory: Washington, DC, 2002.

Perzanowski, D.; Brock, D.; Adams, W.; Bugajska, M.; Schultz, A.; Trafton, J.; Blisard, S.; Skubic, M. *Finding the FOO: A Pilot Study for a Multimodal Interface;* ADA434963; U.S. Navy Center for Applied Research in Artificial Intelligence, U.S. Naval Research Laboratory: Washington, DC, 2003.

Pettitt, R.; Redden, E.; Carstens, C. *Scalability of Robotic Controllers: An Evaluation of Controller Options;* ARL-TR-4457; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2008.

Prasov, Z. Shared Gaze in Remote Spoken HRI During Distributed Military Operations. *Proceedings of the 7th ACM/IEEE International Conference on Human Robot Interaction*, Boston, MA, 5–8 March 2012.

Redden, E.; Elliott, L. R. Robotic Control Systems for Dismounted Soldiers. In *Human-Robot Interactions in Future Military Operations;* Barnes, M. J., Jentsch, F. G., Eds.; Ashgate Publishing: Farnham, UK, 2010; pp 335–352.

Redden, E.; Carstens, C.; Pettitt, R. *Intuitive Speech-Based Robotic Control;* ARL-TR-5175; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2010.

Schuster, D.; Jentsch, F.; Fincannon, T.; Ososky, S. The Impact of Type and Level of Automation on Situation Awareness and Performance in Human Robot Interaction. In *Engineering Psychology and Cognitive Ergonomics: Understanding Human Cognition;* Springer-Verlag: Berlin, 2013; pp 252–260.

Urban, M.; Bajcsy, P. Fusion of Voice, Gesture, and Human-Computer Interface Controls for Remotely Operated Robot. *Proceedings of the 8th International Conference on Information Fusion*, Philadelphia, PA, 25–28 July 2005.

Appendix. Demographics

This appendix appears in its original form without editorial change.

DEMOGRAPHICS

SAMPLE SIZE = 21

MOS		RANK	AGE	DUTY POSITION
11B – 5	12V – 1	E2 – 3	24 years	Bridge Crew Member – 3
12C – 4	12W – 2	E3 – 7	(mean)	Carpentry/Mason Spec – 2
12N – 5	18X – 1	E4 – 6		Crew Chief – 1
12R – 1	29E – 1	E5 – 5		EW NCOIC – 1
12T – 1				Engineer - 1
				Heavy Vehicle Driver – 1
				S1 NCOIC – 1
				Team Leader – 2
				NR – 9

1. How long have you served in the military? 32 months (mean)

2. How long have you been deployed overseas? 22 months (mean)

3. How long have you been deployed in a combat area? 19 months (mean)

4. What is corrected visual acuity? 16 20/20 both eyes 2 20/20 one eye 2 Other

5. Please list any visual problems you have. 1 Astigmatism 1 Color blindness
2 Other

6. What is your height? 69 inches (mean) – range is 63 to 78 inches

7. What is your weight? 172 pounds (mean) – range is 127 to 218 pounds

8. With which hand do you most often write? 20 Right 1 Left

9. With which hand do you most often fire a weapon? 19 Right 2 Left

10. Do you wear prescription lenses? 5 Glasses 2 Contacts 1 Both

11. Do you wear prescription lenses while firing a weapon? 11 No 6 Yes 4 NR

12. Which is your dominant eye? 17 Right 3 Left 1 NR

13. Have you ever driven a robotic vehicle? 17 No 4 Yes

	Responses
While deployed, IRobot/Talon.	1
Throwbot.	1

14. Have you ever driven a remote control car? _6_ No _15_ Yes

15. Have you ever driven a robot on a video game? _10_ No _11_ Yes

An EOD robot to disable mines.	2
EOD, similar to a remote control car.	1
Black Ops II	1
COB Black Ops exploding vehicle.	1
Call of duty.	3
A little bit difficult to control at first.	1

16. Have you ever used speech/voice activated control to operate a device?
_16_ No _5_ Yes

Phone.	3
Computer.	1
Annoying; sounds in the background often makes them do other tasks.	1

17. Is English your second language? _20_ No _1_ Yes

Spanish.	1

18. Where did you spend most of your time growing up (i.e., Georgia, Ontario, etc.)?

US. Military brat; moved all over.	1
Minnesota.	1
Michigan.	1
Colorado	1
Alabama	1
Texas.	1
Texas/Arkansas	1
California.	2
Illinois	1
Ohio	1
North Carolina	2
Georgia	1
New York	1
Florida	1
Canada	1
Peru.	1
Guam	1

19. Has anyone ever told you that you had a regional accent? __9__ No __12__ Yes

	Responses
Northern.	2
Yankee.	1
Southern.	5
Texan	1
Hispanic.	1
Islander	1

20. Using the scale below, rate your level of experience with the following computer software and computer-related activities.

1	2	3	4	5	6	7
No experience	Below average	Slightly below average	Average	Slightly above average	Above average	Expert

	MEAN
Microsoft Windows 98, 2000, XP, etc.	4.86
Computer-based games	4.95
Army digital systems (e.g., FBCB2)	3.38
I would self-rate my computer skills as:	4.67

21. Using the scale below, please self-rate the following Knowledge, Skills, and Abilities (KSAs) related to military duties.

1	2	3	4	5	6	7
No experience	Below average	Slightly below average	Average	Slightly above average	Above average	Expert

	MEAN
Knowledge of tactics, techniques, and procedures (TTP)	4.33
Knowledge of map reading and orientation in field setting	4.33
Knowledge of reconnaissance, surveillance, and target acquisition procedures	3.95
Knowledge relating to communications equipment & communications procedures	4.10
Communication skills (ability to use communications equipment and face-to-face communications to enhance mission accomplishment)	4.43

POST ITERATION

SAMPLE SIZE = 21

TRAINING

1. Using the scale below, please rate the training you received in the following areas.

1	2	3	4	5	6	7
Extremely bad	Very bad	Bad	Neutral	Good	Very good	Extremely good

	Speech		Manual	
	Mean	***SD***	**Mean**	***SD***
a. Explanation of how to drive robot	6.6	*0.9*	6.5	*0.7*
b. Time provided for explanation of the controller	6.6	*0.8*	6.4	*0.8*
c. Training on how to use the controller	6.6	*0.7*	6.4	*0.7*
d. Explanation of how to complete each station of course	6.7	*0.7*	6.4	*0.8*
e. Amount of practice time	6.4	*0.9*	6.3	*0.8*
f. Overall evaluation of the training	6.7	*0.6*	6.6	*0.6*

Comments	**No. of Responses**
Speech	
They were patient on explaining and waited until I was comfortable.	2
Very good initial instruction.	1
I feel the amount of time used to explain the device was very well used and explained well.	1
Good training with voice command.	1
Clear and concise.	2
Easy to understand.	2
Thorough and simple.	2
Robot had some issues understanding my accent.	1
Sometimes you need to repeat your commands.	1
You should inform tester than the backpack gives off a lot of heat.	1
Manual	
The robot was very easy to control and operate.	1
It was good training.	1
Trainers were helpful and informative.	1
When explaining how to use the controller, make sure to explain that wherever the mouse is clicked is where the robot will go vs. just the area of the grid letter that its pointed in.	1
More training time to get used to the controller.	1

2. What were the easiest and hardest training tasks to learn?

1	2	3	4	5	6	7
Extremely hard	Very hard	Hard	Neutral	Easy	Very easy	Extremely easy

	Speech		Manual	
	Mean	***SD***	**Mean**	***SD***
a. Directing the robot	5.6	*1.0*	5.7	*1.0*
b. Operating the controller	6.1	*0.9*	5.6	*1.0*
c. Using the controller in the GRID mode	6.4	*0.7*	6.0	*1.0*
d. Using the controller in the OS mode	4.9	*1.9*	5.4	*1.5*

Comments	**No. of Responses**
Speech	
Exceptional piece of equipment; easy to understand.	1
Better sense of direction with the GRID mode.	1
With GRID you can tell it where to go.	1
The GRID mode is very easy to use for an inexperienced operator. Personally prefer GRID mode. Field of view stays clearer.	1
d. OS mode voice navigation: Select "letter" not being set and randomness makes route difficult.	1
OS mode was harder because you had parts you could only let it go too. If you don't speak loud enough it will continue on the last given command.	1
OS mode is difficult due to the constant change of points on the controller. It's much more complicated to use command.	1
With OS mode the item numbers change while moving causing slight distraction.	1
OS was difficult because the objects letter identifying them kept changing.	1
In OS, difficult to make on-the-fly corrections due to constant updating of tracking. Recommend integrated turn left and right or use a degree-based turning system.	1
Robot does not work respond well in OS mode.	1
The robot does not respond to command "stop" very well at times. I had to say it 3-4 times.	1
Manual	
Training on operating the controller and robot was easy and dummy proof.	1
The time delay between command and execution makes direction slightly difficult. OS mode is kind of distracting. Makes identifying multiple objects difficult. Some blur together with shadows.	1
The problem with OS mode was the amount of objects the robot picked up was making it more challenging to operate.	1
It's really hard to try to detect UXOs while on OS mode because the lines on the screen and delay. It's also hard to make it stop where it needs to.	1
Sometimes when choosing a desired position on the GRID, the robot would	1

Comments	No. of Responses
go off course or outside the engineer tape.	
While directing the robot, it would sometime not follow the path of direction given by operator. The robot would not always stop when given the command.	1
You weren't able to adjust the course of the robot with the left and right buttons.	1
Hardest thing was seemingly stopping the robot even while on the GRID. I had to go to the cancel option a lot.	1
If cursor wasn't moved far enough or the cursor wasn't in a readable place, the stop button wouldn't work.	1
The delay in video to the actual place of the robot.	1

ROBOTIC COURSES

1. Using the scale below, please rate your ability to perform each of the following <u>control tasks</u> based on your experience with the controller you used.

1	2	3	4	5	6	7
Extremely difficult	Very difficult	Difficult	Neutral	Easy	Very easy	Extremely easy

	Speech		Manual	
	Mean	***SD***	**Mean**	***SD***
a. Staying within course limits (engineer tape)	5.2	*1.3*	5.4	*1.1*
b. Avoiding UXOs	5.0	*1.4*	4.8	*1.4*
c. Observing UXOs	5.3	*1.3*	5.1	*1.3*
d. Directing the robot using the OS mode	4.1	*1.9*	5.1	*1.6*
e. Directing the robot using the GRID mode	6.2	*0.8*	6.0	*1.0*
f. The systems responsiveness to your commands (manual inputs/speech)	4.5	*1.3*	5.3	*1.3*
g. Ease of choosing the desired GRID	6.0	*1.0*	5.8	*1.0*
h. Ease of choosing the desired object in OS mode	4.1	*2.2*	5.3	*1.3*
i. Maintaining Situation awareness of items on the course	5.3	*1.3*	5.3	*1.1*
j. Maintaining situation awareness around you	5.2	*1.4*	5.0	*1.4*
k. Ease of maneuvering between waypoints while performing secondary tasks	5.4	*1.1*	4.9	*1.3*
l. Ease of maneuver between waypoints with no secondary tasks	6.2	*1.0*	5.9	*0.8*
m. Maintain control of robot while negotiating the course	5.3	*1.1*	5.2	*1.1*

Comments	No. of Responses
Speech	
Speech mode was easier than manual due to the quicker/faster response time of robot obeying commands.	1
Ease of choosing the desired grid.	1
Maintaining control of the weapon was extremely easy using the voice command and made secondary tasks effortless.	1
Voice over using control was better for the fact I had my hand on the weapon more and had better awareness around me.	1
Grids covered a wide area of section. Good for long hauls but bad for a more direct target.	1
I had not problems avoiding UXOs which was an issue using the manual condition. However, maintaining situation awareness around the robot lacks the peripheral view still.	1
The monitor is delayed and makes it difficult to tell the robot when to stop.	1
Screen resolution still slight factor in difficulty.	1
Robot is always running into objects making it difficult to pay attention to where the robot is going and look for UXOs at the same time.	
Slight difficulty in voice recognition.	1
Robot had hard time understanding my accent.	1
I had trouble telling it to stop when needed it too.	1
Robot responds better in GRID mode.	1
h. Due to outlines being unclear.	1
j. Constant need to redirect course requires attention.	1
The points in OS mode were very difficult to see due to glare on the monitor.	1
OS mode sucked. It takes too much time to register what object is going to be what letter and often started to confuse me because the object would change letters really fast.	1
OS mode was difficult to navigate with because the letter would change unless it was completely stopped. Time limit suggests speed, but having to stop to accurately choose object would hinder that process.	1
With the course in place, a controlled environment as such, I would still like to see the equipment operate in a rougher environment, i.e., off of paved and smooth surfaces. How well would it operate in OS mode going over rough terrain?	
Manual	
a-h: Screen resolution low; could've been lighting.	1
i-j: During avoidance course, more attention was required to avoid UXOs and other objects. Light: uncontrollable.	1
m: Time lag.	1
Robot responds well to manual commands.	1
The UXOs were sometimes blurry.	1
Sometimes when you choose a point it goes around the less desired route.	1
The most difficult part of the task is in OS mode. As the robot moves it picks up small objects to large objects in the distance bunching together. The	1

Comments	No. of Responses
response time with the robot is difficult to stay within the course due to a lag after commands. It seems you really have to speak as clear as possible in a commander's voice to get a response.	
In OS mode the robot picked up items that were not relevant to navigation (pine needles, grass, spots on the ground, etc.). Also the robot wouldn't lock onto an object like in GRID mode.	1
It was difficult to keep my firing hand on the weapon while trying to control the robot.	1
I did not like the fact that I had to take my hand off my weapon so much to maneuver it. Also, I do not like having to look at my arm. I lose too much awareness. You need to have a heads-up display.	1
It's hard for robot to stop on command. Robot missed lots of small UXOs. OS mode is difficult with lines all over screen. Robot failed to respond to manual commands.	1
In OS mode, when you selected an object farther away, seemed more difficult.	1
Avoiding UXOs was difficult due to the robot not being familiar with it and not adjusting itself to move around the UXO. Also, I would command the robot to stop itself and sometimes it failed to do so, resulting in hitting the UXO. Since the UXOs are small in size, it's difficult to observe or know if it's an UXO or not. Maintaining situational awareness to its fullest capacity was not valid due to its peripheral vision being obstructed.	1
When the auto avoidance kicked in, the stop button wouldn't respond or had a long delay.	1
Left/right and cancel button too small on the display.	1

2. During this trial, please check any of the following problems you may have experienced?

	Number of Responses	
	Speech	**Manual**
Wrist strain	0	3
Arm strain	3	4
Eye Strain	2	2
Hand strain	0	0
Finger strain	0	0
Nausea	0	0
Fatigue	0	0
Disorientation	0	0
Dizziness	0	0
Other	0	1

Comments	No. of Responses
Speech	
Very little arm strain.	1
Hard to bend arm with control box arm.	1
Slight eye strain.	2
Manual	
Slight eye strain.	1
A little wrist strain.	1
Cannot bend arm with deuce on.	1
Neck strain.	1
Glare on screen.	1

3.a Using the scale below, what is your overall rating of the control method (speech/ manual) that you used?

1	2	3	4	5	6	7
Extremely bad	Very bad	Bad	Neutral	Good	Very good	Extremely good

Speech		Manual	
Mean	***SD***	**Mean**	***SD***
5.9	*1.1*	5.5	*1.2*

b. What are your comments on the control method (speech/manual)?

Comments	No. of Responses
Speech	
All control methods were very easy to learn and use.	3
Very cool. Fun and interesting.	1
Speech was overall better than manual. Had more control that was faster at hand.	3
Highly effective compared to the manual condition. It was simple maneuvering and effective considering the mental aspect given.	1
I didn't expect it to pick up every command especially a person with a slight accent. A note I had to make to myself was stopping the robot then repeat command.	1
Very good.	2
It was good. I just need to be louder and clearer on commands given.	1
Overall, the robot seemed effective in GRID mode even with distractions and obstacles.	1
Really like the speech GRID mode. Maintains clear sigh as travel occurs. Very easy to operate. User friendly.	1
Simple enough. Sometimes I would have to repeat myself.	1
GRID superior.	1
Works well except in OS mode as it does not seem to respond to the	1

Comments	No. of Responses
commands very well.	
Occasional miscommunication with voice recognition.	1
Had to repeat my commands.	1
Maybe a slight combination of both modes?	1
Only problem was calling “spin right.”	1
Sometimes the robot would start going in circles.	1
Instead of saying “spin” it is more natural for people to say “turn.” And the command “stop” usually didn’t work.	1
OS was difficult.	1
OS mode could be corrected if it picked up large distinct objects with better response time.	1
OS was difficult only in an object filled area.	1
I think it will be better to direct the robot by numbers and distance. Like 12 o’clock 15m.	1
Only thing bad was in GRID. It goes to an area target, while OS will highlight entire target and it will take you directly to desired target.	1
Manual	
The manual is a great way to move the robot when you have to be silent and hidden. The reaction could be faster, but other than that it was great.	1
The manual control method was very good because the viewpoint was adequate and it automatically adjusted itself to move around objects.	1
Very easy to use.	2
I found it very easy to control the robot.	2
Attainable, easy to learn information on how to use the system. Manual input seemed slower, mostly through controller error.	1
Object recognition and avoidance great.	1
After the first few minutes of use, controlling the robot became second nature.	1
A manual reverse button.	1
A screen overlay to counter glare.	1
It seems to take longer to use the mouse to choose your desired location.	1
Was a bit slower and I had my hand off my weapon longer but had better control.	1
Manual mode feels like it would be used while sitting in a secure location far from the IED field (i.e., HMMWV, Buffalo, etc.).	1
Robot could be more accurate when going to selected point.	1
I think the screen should be to control. Tap to move, double tap to stop (would be great if we can zoom in/out with finger like an IPOD).	1
The stop button did not always work.	1
When the robot came up to a few close objects I feel like it did not know what to do with the self correction.	1

4.a. Using the scale below, what is your overall rating of the OS mode that you used?

1	2	3	4	5	6	7
Extremely bad	Very bad	Bad	Neutral	Good	Very good	Extremely good

Speech		**Manual**	
Mean	***SD***	**Mean**	***SD***
4.3	*1.8*	5.0	*1.5*

b. What are your comments on the OS mode that you used?

Comments	**No. of Responses**
Speech	
Works very well with identifying objects; however it changes object identifications while driving causing slight view obstruction.	1
OS mode seemed to be a faster way to navigate through the course. I just had to be aware of the grid identifiers changing constantly.	1
OS mode made it easy to get closer to a specific target but I preferred GRID mode to go between two objects.	1
OS mode is extremely better than GRID. OS takes you to the target but doesn't always highlight all objects to you don't always have an option.	1
OS mode is simple, but the downside is finding the right obstacles to use as a way point and waiting for the letters to load.	1
I like the idea of object recognition. Still needs work, however.	1
Recommend better on-the-fly programming along with degree-based turning system.	1
It's a little harder going from object to object. The camera has to be still for the letters to appear again. It's got about a 3-4 second lag which makes it difficult to use.	1
It did not mark the UXOs, but the USOs were mostly within clear picture. One problem occurred when the robot detected grass. It lost the command given and started to move off course; however, I immediately stopped the robot and turned it to the corrective action.	1
Very hard to keep the letter on object and then wait to reappear when robot stops.	1
OS mode did not respond very well to the voice commands. The letters were very difficult to see. The letters continued to move which made it very hard to pick a point.	1
OS mode picks up everything from rocks to piles of pine needles. When it picks up larger objects like silhouettes, it puts 4-8 different phonetic options on the screen for one object and it constantly changes your options make it difficult to choose.	1
In an object rich environment, the OS mode switches letters on objects in the middle of movement making it hard to redirect the robot.	1
OS objects change too quickly. I would select one and by the time I said it,	1

Comments	No. of Responses
the objects changed and set the robot into a different direction.	
OS took way, way too long to pick out objects. It also made the screen too hard to see objects.	1
OS points move around too much.	1
The only bad thing is that even after you stop, the targets keep changing a few more times.	1
I liked it less than GRID because after giving it a command, I couldn't control it until it came to a stop because the call signs would change until it did.	1
Manual	
Good experience.	1
Manual works best for OS mode.	1
Very good for on-the-fly corrections.	1
OS mode is great if there are many objectless.	1
Easier to identify objects/targets outlines moving across screen. Very small distraction.	1
Works much better in manual mode than speech mode.	1
Better than GRID because it can move with less stop and go.	1
The picture was clearer compared to the GRID mode.	1
I was able to complete the course a lot faster, and without as many errors.	1
This time I could tell it exactly where I wanted it to go with no mishaps.	1
I enjoyed the physical multitask work and especially the running. It tested my skills of being multitasked.	1
I like the concept. If refined to differentiate between singular objects would make identification easier. Especially in low lighting.	1
Easy to use, but I felt more comfortable in GRID mode changing direction while the robot was in motion.	1
It wasn't bad.	1
It didn't do much.	1
Not as accurate moving to the path of direction given.	1
Still challenging to go around objects without it hitting something.	1
It jumbles up the screen often changing phonetic selections often getting you to change course making it inefficient.	1
Lines on screen make it difficult to get a good view of terrain.	1
With the screen delaying and robot moving, we missed a section of the route.	1
You had to click on an object to get the robot to move. Some time the objects were too large and it would send the robot in a general direction. The tree line would be on large object.	1
Only problem was controller error with getting the robot to stop.	1
OS mode points change too rapidly.	1

5.a. Using the scale below, what is your overall rating of the GRID mode that you used?

1	2	3	4	5	6	7
Extremely bad	Very bad	Bad	Neutral	Good	Very good	Extremely good

Speech		Manual	
Mean	***SD***	**Mean**	***SD***
6.3	*0.7*	5.9	*1.1*

b. What are your comments on the GRID mode that you used?

Comments	**No. of Responses**
Speech	
Excellent.	1
Easy task.	1
GRID is easier.	2
Effective.	1
Very user friendly.	2
Very easy to navigate with no problems.	1
Faster and better than OS mode. Did not have to wait for OS to pick out objects. Say and go, on the go.	1
GRID mode made it very stress free to move/control the robot.	1
GRID mode was dead on. Didn't have any major issues. A little more meticulous and slow approach, but overall well.	1
Very easy to control the robot and choose the best course path. I felt much more confident controlling the robot in the GRID during any situation.	1
My personal choice. More precise than OS.	1
Very good system. Still recognized all objects.	1
Overall good but does stop short of target. I like the obstacle avoidance on it.	1
The GRID was simple and I could change the course while moving easily and know where it was headed.	1
GRID is easier because you don't have to stop giving it commands. It will turn in the middle of a command if you need it too. It's easier than stop-go, stop-go.	1
Not having outlines (like OS) made for a clearer view of operations.	1
Manual	
Very easy to use.	2
Simple to learn.	1
Personally I like this more. Very clear field of view.	1
A lot easier, mainly because I knew the grid point wasn't changing so often.	1
I enjoyed the GRID mode due to the squares on it helping myself marking the cursor where the robot would move itself towards. It seemed more accurate and it was easy to navigate. Great experience.	1
Very clear visual space to maneuver the robot through the course; goes generally toward selected path desired.	1

Comments	No. of Responses
It was easier to just choose a spot to go.	2
The GRID is great because you can select a set distance and point or make the robot move wherever you want it to go.	1
Starting and getting the robot to a certain spot was easier. You could click on the horizon and make minor adjustments.	1
With less obstruction on the screen than the incomplete OS mode, navigation was easier.	1
It was harder to pick a point with the GRID mode vs. the OS mode where an area was outlined for you to click on.	1
I felt I could not see as much with the grids on the screen.	1
It didn't do much.	1
Too much stop and go. Better to have the OS mode. Just click and go.	1

6. Using the scale below, please rate the response to SPEECH commands

1	2	3	4	5	6	7
Totally disagree	Slightly disagree	Disagree	Neutral	Agree	Slightly agree	Totally agree

	Mean	***SD***
a. The robot responded well to "Select" commands	5.6	*1.0*
b. The robot responded well to "Spin" commands	5.0	*1.4*
c. The robot responded well to "Stop" commands	4.9	*1.8*
d. The robot responded well to all my commands	5.1	*1.1*

Speech

I had to repeat a couple of commands, spinning right/left. 1

7. General Comments:

Speech

My husband wants to go EOD and if this is the robot they're going to use, I'm not letting him do it. 1

END OF EXPERIMENT

SAMPLE SIZE = 21

1. Please select the control condition you felt was the most effective when performing the following tasks (you can check more than one condition for each task):

TASKS	Number of Responses				
	Manual OS	**Manual GRID**	**Speech-based OS**	**Speech-based GRID**	**None**
Staying within course limits (engineer tape)	4	10	5	15	-
Maneuvering the robot to a desired location	4	14	3	14	-
The systems responsiveness to your commands	7	16	3	9	-
Maintaining situation awareness of items on the course	6	11	7	16	-
Maintaining situation awareness around you	4	8	7	17	-
Ability to perform secondary tasks	3	5	10	16	-
Maintain driving control while negotiating the course	4	15	5	13	-
Overall Task performance	3	11	5	11	-
List any additional task you may want to add below:					
--Turn around	-	1	-	1	-
--Maneuver course while hands-free	-	-	-	1	-

Comments	**No. of Responses**
Both GRID and OS for speech were a lot easier and it allowed me to keep my hand on my weapon.	1
Speech-based GRID is a very good system; easy to use; and very reactive.	1
Overall, the Speech-based GRID prevailed over its other conditions.	1
Either grid mode works well. There is a lot of visual room to see upcoming objects.	1
OS has potential. Have it recognize solid planes, outline, and then disregard them in favor of other objects.	1
In OS mode the screen seems to jumble up with unnecessary objects being highlighted and the phonetic selections constantly changing making it more difficult to complete tasks.	1
I prefer GRID to OS because it has more options and doesn't require objects to be present.	1
Better to use numbers 1 to 12 and distance instead of the alphabet because people like me and the strong accent.	1
No additional tasks really needed.	1

2. Do you have suggestions for ways to increase the effectiveness of the **MANUAL** mode?

14 Yes 7 No

Comments	**No. of Responses**
Works a lot better than speech. It feels as though you have more control.	1
Place buttons on side of screen for turn and stop.	1
Increase size of the spin left/right/cancel buttons or add them to the buttons.	1
Bigger curser and icons.	1
Have the controls built into the screen.	1
Mount control on weapon.	1
Make screen touch sensitive with a stylus and a glare cover.	1
Improve the stop command on the technology device.	1
When you click on an object, it would be better if the robot targeted that object instead of going in a general direction.	1
Touch screen.	1
I'm used to using a touchscreen so that might have helped me more than the mouse.	1
Touch screen or making one of the buttons on the mouse a designated stop button. Maybe make the center button a joystick?	1
Maybe system response on stop. Just keep right button as stop no matter if it's on the grid or not.	1
Zoom in/out option.	1
The control is a waste of space (get rid of it).	1
Try to put the controller built onto the screen so there isn't any strain on wrist. It's all held up by forearm.	1
If on arm, need touchscreen; faster. However, it needs to have a HUD system. Thus, a person can look around and see what is going on around him/her, i.e., an eye system.	1

3. Do you have suggestions for ways to increase the effectiveness of the **SPEECH** mode?

9 Yes 11 No 1 NR

Comments	**No. of Responses**
Improve recognition software, obviously.	1
Try putting a more precise command on there.	1
More simple commands like slight right/left while driving.	1
Go forward, turn not spin, go backward, spin 160, 360.	1
Make the robot be less decisive when you call out a command, such as stop. Such a simple command that it often failed to follow.	1
Numbers and distance instead of alphabet I think will work better.	1

Comments	**No. of Responses**
Make it so it responds when you give it a command the first time.	1
Need to speak over and over.	1
Have a mini-size mic attached to shirt.	1
Wireless headset.	1
Larger display screen.	1

4. Do you think that a Speech-based controller shows potential for robotic operation?

_20_ Yes _1_ No

Awesome technology.	1
Very great system.	1
Highly effective during battle because it's extremely fast and easy.	1
Responsive to commands.	1
It is faster and simpler for someone to control a robot.	1
Faster and you have awareness around you.	1
Makes multitasking simple; maintaining situational awareness.	2
Allows you to operate the robot without using your hands.	1
Hands-free will improve commands on the robot while soldier keeps hands for weapon, radio or whatever the case may be.	1
Soldiers will not always have time to manually direct said robot.	1
Robots with speech controlled elements would be a great boost to units overseas that also need to keep hands-free.	1
Just as effective as manual mode.	1

5. Do you think that a **GRID-**based controller shows potential for robotic operation?

_21_ Yes _0_ No

Easy to operate.	2
Allows for clearer vision.	1
More exact specification of location desired to move to.	1
By far the best operational mode for this equipment, manual/speech.	1
In light conditions it is perfect.	1
It is a precise means to direct a robot.	1
I believe the GRID-based controlled is better and more effective at giving commands and easier to navigate.	1
Allows you to control the robot more effectively.	1
Would work better when no noise was allowed, and when there is a lot of background noise.	1
When you cannot speak, go GRID.	1

6. Do you think that a **OS-**based controller shows potential for robotic operation?

_15_ Yes _6_ No

Comments	**No. of Responses**
Good system.	1
Yes, for manual operations, the speech OS would just need to not switch letters during movement and be a great asset.	1
In the dark it seems better to use. You can see everything in the day, but when finding objects in the dark it is really useful.	1
Can be used to see objects and verify that they're there.	1
Maybe if the bugs are worked out better, I would change my opinion.	1
There is a lot to improve here, but it does have a bright future.	1
However, not on its own. Coupled with the grid mode would allow for better object recognition and avoidance.	1
It has trouble identifying actual objects of interest. Changing its selection often while moving through the course. The phonetic selection on screen changes by the second, making it very hard to call out commands.	1
Can only move to recognized objects. Outlines slightly obstructive of view.	1
It causes too m any distractions in the screen. Slow and delays a lot.	1
Takes too long to pick out objects.	1
Only in Manual because of on-the-fly corrections, but if you can integrate course corrections on-the-fly for Speech, it would be extremely effective.	1
Points just move too much.	1

7. Using the scale below, please rate the following characteristics of the controller system.

1	2	3	4	5	6	7
Extremely bad	Very bad	Bad	Neutral	Good	Very good	Extremely good

CHARACTERISTICS	**Mean**	***SD***
Size of the controller	5.1	*1.3*
Size of the display screen	5.6	*1.1*
Clarity of the display screen in bright sunlight	4.9	*1.4*
Clarity of the display screen in the shade	5.7	*1.1*
Placement of the display screen on the forearm	5.5	*1.2*
Weight of the controller	5.5	*1.4*
Weight of the control system (backpack, display, mouse)	5.4	*1.0*
Color of the grid pattern	5.6	*1.0*
Color of the OS pattern	5.2	*1.5*
Comfort of the controller	5.4	*1.2*
Obstacle Avoidance capability	5.3	*1.1*

Comments	No. of Responses
Screen size was good.	1
Weight was fine. Any additional gear would be a problem.	1
Controller felt bulky and a little unwieldy. Maybe slim it down?	1
The screen rotated on my forearm making the glare effect come into factor while trying to maneuver the robot while performing secondary tasks.	1
Needs improvement with small objects (obstacle avoidance).	1
This system (manual and speech) needs an option to turn robot 45^o, 90^o, and 180^o at one time rather than having to spin it multiple times to get it there.	1
Color (OS pattern) pattern made screen too hard to see.	1

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

2 (PDF) DIRECTOR
US ARMY RESEARCH LAB
RDRL CIO LL
IMAL HRA MAIL & RECORDS MGMT

1 (PDF) GOVT PRINTG OFC
A MALHOTRA

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM C A DAVISON
320 MANSCEN LOOP STE 115
FORT LEONARD WOOD MO 65473

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM D
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL 35898-7290

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRS EA DR V J RICE
BLDG 4011 RM 217
1750 GREELEY RD
FORT SAM HOUSTON TX 78234-5002

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM DG J RUBINSTEIN
BLDG 333
PICATINNY ARSENAL NJ 07806-5000

1 (PDF) ARMY RSCH LABORATORY – HRED
ARMC FIELD ELEMENT
RDRL HRM CH C BURNS
THIRD AVE BLDG 1467B RM 336
FORT KNOX KY 40121

1 (PDF) ARMY RSCH LABORATORY – HRED
AWC FIELD ELEMENT
RDRL HRM DJ D DURBIN
BLDG 4506 (DCD) RM 107
FORT RUCKER AL 36362-5000

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM CK J REINHART
10125 KINGMAN RD BLDG 317
FORT BELVOIR VA 22060-5828

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AY M BARNES
2520 HEALY AVE
STE 1172 BLDG 51005
FORT HUACHUCA AZ 85613-7069

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AP D UNGVARSKY
POPE HALL BLDG 470
BCBL 806 HARRISON DR
FORT LEAVENWORTH KS 66027-2302

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AT J CHEN
12423 RESEARCH PKWY
ORLANDO FL 32826-3276

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AT C KORTENHAUS
12350 RESEARCH PKWY
ORLANDO FL 32826-3276

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM CU B LUTAS-SPENCER
6501 E 11 MILE RD MS 284
BLDG 200A 2ND FL RM 2104
WARREN MI 48397-5000

1 (PDF) ARMY RSCH LABORATORY – HRED
FIRES CTR OF EXCELLENCE
FIELD ELEMENT
RDRL HRM AF C HERNANDEZ
3040 NW AUSTIN RD RM 221
FORT SILL OK 73503-9043

1 (PDF) ARMY RSCH LABORATORY – HRED
RDRL HRM AV W CULBERTSON
91012 STATION AVE
FORT HOOD TX 76544-5073

1 (PDF) ARMY RSCH LABORATORY – HRED
HUMAN RSRCH AND ENGRNG
DIRCTRT MCOE FIELD ELEMENT
RDRL HRM DW C CARSTENS
6450 WAY ST
BLDG 2839 RM 310
FORT BENNING GA 31905-5400

1 ARMY RSCH LABORATORY – HRED
(PDF) RDRL HRM DE A MARES
1733 PLEASONTON RD BOX 3
FORT BLISS TX 79916-6816

8 ARMY RSCH LABORATORY – HRED
(PDF) SIMULATION & TRAINING
TECHNOLOGY CENTER
RDRL HRT COL M CLARKE
RDRL HRT I MARTINEZ
RDRL HRT T R SOTTILARE
RDRL HRT B N FINKELSTEIN
RDRL HRT G A RODRIGUEZ
RDRL HRT I J HART
RDRL HRT M C METEVIER
RDRL HRT S B PETTIT
12423 RESEARCH PARKWAY
ORLANDO FL 32826

1 ARMY RSCH LABORATORY – HRED
(PDF) HQ USASOC
RDRL HRM CN R SPENCER
BLDG E2929 DESERT STORM DRIVE
FORT BRAGG NC 28310

1 ARMY G1
(PDF) DAPE MR B KNAPP
300 ARMY PENTAGON RM 2C489
WASHINGTON DC 20310-0300

13 DIR USARL
(PDF) RDRL HR
L ALLENDER
P FRANASZCZUK
C COSENZO
RDRL HRM
P SAVAGE-KNEPSHIELD
RDRL HRM AL
C PAULILLO
RDRL HRM B
J GRYNOVICKI
RDRL HRM C
L GARRETT
RDRL HRM DW
R PETTITT
RDRL HRS
J LOCKETT
RDRL HRS B
M LAFIANDRA
RDRL HRS C
K MCDOWELL
RDRL HRS D
B AMREIN
RDRL HRS E
D HEADLEY